

Research on Classification of Hyperspectral Remote Sensing Imagery Based on BDT-SMO and Combined Features

Fenghua Huang^{1,2} and YAN Lu-Ming¹

1. College of Geographical Sciences, Fujian Normal University, Fuzhou, Fujian, China

2. Sunshine College, Fuzhou University, Fuzhou, Fujian, China

Abstract—There are some prevalent problems in the classification of hyperspectral remote sensing imagery currently, such as many bands, large amount of data, high proportion of mixed pixels and lower spatial resolution and so on. In order to solve the above problems, the sequential minimal optimization (SMO) algorithm is researched, and a supervised classification method based on binary decision tree SMO (BDT-SMO) algorithm and spectrum-texture combined features is proposed to improve the accuracy and efficiency of hyperspectral remote sensing imagery classification. The higher spatial resolution imagery (ALI) and hyperspectral imagery (Hyperion) which have been acquired from the same sensor (EO-1) in the same time are used as the experimental data and implemented geometric correction and pixel-level fusion. Extract the spectral features and textural features of the ground objects from the fusion images, combine the features above and train the BDT-SMO multi-class classifier based on separation degree. The classifier is used for the land-use status classification of experimental areas. Select two different sets of samples which are based on spectral features and spectrum-texture combined features, and use the four different methods of maximum likelihood, BP neural network, BDT-SVM and BDT-SMO to train the two different sets of samples above separately. Experimental results show that the BDT-SMO classification method based on combined features can improve the efficiency and accuracy of land-use status classification effectively and has better generalization ability.

Index Terms—Hyperspectral Imagery, SMO Optimization Algorithm, Binary Decision Tree, Textural Features

I. INTRODUCTION

The hyperspectral remote sensing images always contain many continuous and narrow wave bands, and the continuous spectral feature curves of the observed ground objects can be formed which have the better ability of ground objects recognition. However, hyperspectral remote sensing image has many common problems such as large amount of data, the higher proportion of mixed pixels, and low spatial resolution, thus the improvement of classification speed and accuracy will be limited if ground objects are classified simply using hyperspectral images. Currently, remote sensing image classification is mainly based on spectrum and texture (two major basic features), and also top priority is given to spectrum. Texture information can allow people to observe the

detailed changes of the surface within a small space scope, and also can reflect the spatial structure and internal difference of different targeted ground objects, so it has been used as an important foundation of satellite image interpretation and information extraction [1].

Jiguang Hong proposed the gray-gradient co-occurrence matrix on the basis of gray co-occurrence matrix in 1984 [2], and also applied 15 features of the model to classifying and recognizing five types of white blood cell samples. His experimental result shows that this method can play a better classification result than that of the gray co-occurrence matrix for the images which have unclear cell nucleus boundaries [2]. Gray-gradient co-occurrence matrix model intensively reflects two types of the most basic information of images (the gray-level and gradient (or edge) of pixels) and their mutual relationship [3]. The gray level of each pixel is the foundation to form an image, while gradient is the element of image edge contours. Gray-gradient space can clearly show the distribution discipline of pixel gray level and gradient in image, and also reflect the spatial relationship of all pixels with the neighborhood pixels, so that image texture can be well described. The classification accuracy will be greatly improved if the spectral features as well as the corresponding textural features are comprehensively applied to remote sensing image classification.

The traditional support vector machine (SVM) can be summarized as a quadratic programming problem, and the most existing training algorithms for the training problems of the large-scale training sample sets are based on the idea of decomposition iteration currently, so that the learning efficiency of the algorithms is low. The sequential minimal optimization (SMO) algorithm is also a decomposition algorithm, and its workspace only contains two samples, but only two Lagrange multipliers are optimized in each step of iteration. SMO is mainly used for solving two problems: (1) use an analysis method to solve a simple optimization problem; (2) choose the optimized Lagrange multiplier strategy [4]. Although there are more sub problems of quadratic programming in SMO, the overall processing speed is greatly increased. SMO is generally unnecessary to process large matrix, so that there are no special requirements for the storage space and the problems of

SVM large-scale samples training can be solved in stand-alone environment. SMO not only supports higher feature dimensions, but also has better performances in algorithm convergence, training speed, and classification accuracy, so that it has better generalization ability than the traditional SVM [5].

The traditional SVM algorithm is only applicable to classification of two classes, and needs to rely on the specific combination of multiple SVM classifiers to complete multi-class classification. At present, several SVM multi-class classification methods are usually used, such as "one against one" (OAO-SVM), "one against all" (OAA-SVM), directed acyclic graph (DAG-SVM), binary decision tree (BDT-SVM) and error correction coding (ECC-SVM) and so on.

Existing related research result [6] shows that OAO-SVM and OAA-SVM have a rapid classification speed. But if there is a large amount of training samples data, the training efficiency of OAO-SVM and OAA-SVM will be greatly reduced, and the existence of unrecognizable areas will be probable in the training samples. Therefore, they are not applicable to the classification of the hyperspectral images with high dimensions and large samples data; The generalization ability of ECC-SVM is nothing to do with the number of feature dimensions, but the coding rules of ECC-SVM are subjective obviously, and the problems such as determining the length of codes, calculating the minimum hamming distance between codes and seeking the optimal arrangement ways are still necessary to be solved. DAG-SVM model, as a tree structure, has a very rapid classification speed, but is unable to solve the problem of unrecognizable areas. However, if the number of classification types is large, the classification efficiency of DAG-SVM will be not high. The selection of root node relies on experience, and the classification accuracy will be directly affected by the selection of root node. BDT-SVM is similar to DAG-SVM in ideas, but the difference is that two SVM classifiers of $n-1$ are only necessary to build in BDT-SVM, while the internal nodes of n ($n-1/2$) are only necessary. In addition, every node of BDT-SVM has only one parent node, and an internal node can be not "one-to-one" SVM classifier; meanwhile every node of DAG-SVM can has multiple parent nodes, and each internal node must be "one-to-one" SVM classifier. BDT-SVM algorithm has higher classification efficiency, fault tolerance, and generalization ability, and its comprehensive classification effect is better than that of other algorithms. If the traditional multi-class learning strategy of SMO ("one against one") is replaced with the binary decision tree method, the effect of SMO multi-class learning will be improved obviously.

In this paper, the traditional SVM algorithm is replaced with SMO algorithm to improve the learning efficiency of large-scale training data; The higher spatial resolution imagery (ALI) and corresponding hyperspectral imagery (Hyperion) which had been acquired from the same sensor (EO-1) in the same time are used as the experimental data and implemented geometric correction and pixel-level fusion. The

spectral features and textural features were extracted from the fusion image and combined in order to improve the classification accuracy; The land-use status in the experimental area was classified using the multi-class SMO algorithm based on separation degree and Binary Decision Tree (BDT-SMO), so as to improve the accuracy and speed of the fusion image classification.

II. BDT-SMO CLASSIFICATION IMPLEMENTATION PROCESS BASED ON COMBINED FEATURES

The hyperspectral remote sensing image's BDT-SMO classification steps based on combined features are shown in Fig 1.

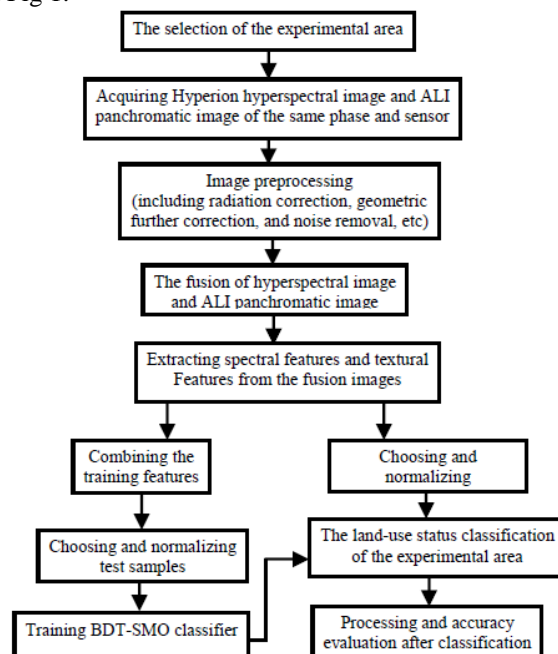


Figure 1. The BDT-SMO classification steps based on combined features

A. Image Preprocessing and Fusion

The experimental area was located near Jinan District of Fuzhou city, and its covering area generally featured parallelogram. The four corners and their locations of the parallelogram are: upper left corner (26.17352500N, 26.17352500E), upper right corner (26.17278333N, 119.33705000E), lower left corner (26.11154167N, 119.29088056E) and lower right corner (26.11080000N, 119.33705000E). Hyperion image (L1G format and 242 bands totally) acquired by EO-1 sensor was used as the hyperspectral data, and the imaging date was on March 26, 2003. The image was clear and included 0% cloud cover. In addition, the spatial resolution of the Hyperion images were 30m, and the study area size was 155×230. EO-1 ALI panchromatic band image of the experimental area was acquired at the same time, and its study area size was 465×690. First, the bands, which were not calibrated and also were greatly affected by water vapor and noise, were removed from Hyperion image, so that 134 bands were finally left. Then the absolute radiation value transformation, bad line repair, stripe removal, Smile effect removal, and Flashed atmospheric correction were

implemented for the above bands, and also the gray levels of pixels were converted to the reflectivity data. Second, radiation calibration, 6S atmospheric correction, and geometry correction were implemented for ALI image. Hyperion image and ALI panchromatic bands were fused using Gramm-Schmidt method, so that the spatial resolution of hyperspectral image was improved. Through the experiment, the spectral feature curves of all ground objects in the images before and after the fusion were qualitatively analyzed. The changes of the image before and after the fusion were also quantitatively analyzed through relevant indicators such as information entropy, average gradient, deviation index, root mean square error, and correlation coefficient.

The experimental result shows that the spatial resolution of the original hyperspectral image was not only improved, but also the original spectral information was well maintained. The hyperspectral image of experimental area after preprocessing and fusion is as shown in Fig 2.



Figure 2. The BDT-SMO classification steps based on combined features

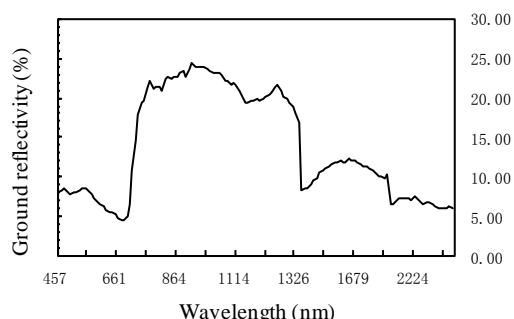


Figure 3. The woodland spectrum feature curve extracted from the pure pixels.

B. Spectrum Feature Extraction

A certain proportion of mixed pixels exist in Hyperion hyperspectral images. Therefore, the interference of the mixed pixels is necessarily eliminated before the image features extraction, so as to extract the relatively pure pixels. Calculate pure pixel index (PPI) By MNF (minimum noise fraction) transform, then specify the

threshold range and get the regions of interest from PPI image. The pixels within the regions of interest are relatively pure pixels. After the spectral feature values (a total of 134 bands) of the pure pixels were extracted, the continuous spectral feature curves will be obtained. The woodland spectrum feature curve extracted from the pure pixels of the fused image is as shown in Fig 3.

C. Texture Feature Extraction

Texture features always contain gray-level distribution, gray gradient (edge), and scale information. Gray-level distribution information mainly reflects the relative position and distribution discipline between the different pixels; gray-level co-occurrence matrix is a very common method used for extracting the gray-level distribution information. Gray gradient information (edge information) mainly reflects the change tendency and boundary features of the gray-level information. For example, roads and water body generally have obvious edge features, and the gray-gradient feature extraction methods mainly include Roberts, Sobel, Prewitt, and other edge extraction algorithms, and they can be extracted through the gray-gradient co-occurrence matrix.

There are 15 common digital features in the gray-gradient co-occurrence matrix [3]: small gradient advantage, large gradient advantage, gray-level distribution non-uniformity, gradient distribution non-uniformity, energy, average gray-level, average gradient, gray-level mean square deviation, gradient mean square deviation, relevance, gray-level entropy, gradient entropy, mixed entropy, inertia moments, and deficit moments. Because the biggest change rate of image gray-level is expressed by gradient, and the density of the gray lines is expressed by the gradient value of every point of the image. Therefore, the intensity of image gray-level changes can be reflected to a certain extent by large gradient advantage and small gradient advantage. If the change of image gray-level becomes gentle, there are more small-gradient pixels, so the advantage of small gradient is greater. If the change of image gray-level becomes intense, there are more large-gradient pixels, so the advantage of large gradient is greater [7].

The textural features of ALI panchromatic band image were extracted using window method, each pixel of the image was corresponding to a window, and 15 gray-level digital feature values sequence of the gray-gradient co-occurrence matrix of every window sub-image were used as the feature vector of the corresponding pixel. A sliding window of $N \times N$ was chosen, and the length of the sliding step was Δx and Δy . The sliding window moved from the top left corner of remote sensing image and then the 15 features values of the gray-gradient co-occurrence matrix of the center pixel of the window were calculated. Then, the above steps were repeated using the sliding window until the whole image area was scanned completely, and the insufficient section of the edges was filled with 0. The different feature values generated were saved into a new band, and finally a total of 15 textural feature bands were generated. The above operations were repeated respectively with different sizes of sliding window and sliding step, and also the results of the operations were

TABLE I. THE SELECTION OF THE SAMPLES IN THE EXPERIMENTAL AREA

Land-use type	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12
The number of Training samples	304	264	419	269	353	308	324	315	302	647	386	306
The number of test samples	254	231	263	251	266	283	278	252	261	272	253	252

compared and analyzed. The results of the analysis show that a sliding window of 15×15 and a sliding step length of 1 ($\Delta x=1, \Delta y=1$) were more appropriate.

D. The Combination of Classified Features

There are a large area of woodland and farmland in the experimental area, and the vegetation characteristics are very obvious. Therefore, the normalized difference vegetation index (NDVI) was used as an important feature to distinguish the different vegetation. Thus, the eventually applied classification features included 134 spectral features, 15 texture features, and NDVI. The feature vector of the classification was constituted by a total of 150 features.

E. Sample Selection and Normalized Processing

According to the latest national land-use status classification standard (GB/T 21010-2007), land-use status is divided into 12 primary types and 127 secondary types. According to the land-use actual status of the experimental area, only 12 primary types were used: Farmland (C1), Woodland (C2), garden land(C3), grassland (C4), the commercial take land(C5), mining and warehouse land (C6), residential land (C7), public administration and public service land (C8), special land (C9), transport land (C10), waters and water conservancy facilities (C11), other lands (C12). Through the field investigation and comprehensive analysis of the remote sensing image, the samples of ground objects were extracted from the representative regions. The number of the training samples should not be less than that of the test samples in general; meanwhile the number of the training samples for each type of ground objects should be not less than that of the classified feature dimensions. In the experimental area, the good representative samples of each land-use type were chosen respectively with the aid of high-resolution images and the 1:10000 land-use status map of the experimental area. The selection of the samples was as shown in table I.

After relevant samples were extracted, the samples must be normalized and formatted.

F. Training BDT-SMO Classifier

In this paper, the traditional SVM was replaced with the SMO algorithm, and hyperspectral images were classified using the binary decision tree SMO (BDT-SMO) based on separation degree, so that a good effect was achieved. The common measure of the separation between different classes was mainly the Euclidean distance between different gravity centers. In Fig 4, it is known that there may be three types of spatial distributions between classes under the same distance of between gravity centers. In sub-image (a), different classes are contained with each other, and they can be directly combined together in the classification; in sub-

image (b), there is no disjoint between different classes, and the classes are easy to separate; in sub-image (c), the classes are partly intersected and difficult to separate. Thus, the difficulty degree of separation is related to not only the distance of gravity centers between different classes but also the radiuses of the circles surrounding the samples of different classes. It was assumed that d_{ij} was the distance between the gravity centers of class i and class j , and the radius r_i of class i was defined as the distance from gravity center c_i of class i to the fastest boundary sample point. The boundary distance between class i and class j was defined as follows:

$$s_{ij} = d_{ij} - (r_i + r_j) \quad (1)$$

The separation degree g_{ij} between class i and class j was defined as follows:

$$g_{ij} = \frac{d_{ij}}{r_i^2 + r_j^2} \quad (2)$$

Class i and class j were more easily separated if g_{ij} was greater.

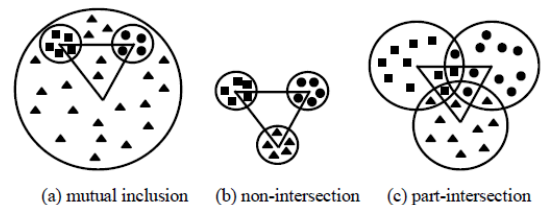


Figure 4. The spatial distribution difference between classes

BDT-SMO training steps based on the degree of separation are as shown below [8]:

The values of boundary distances (s_{ij}) between the sample sets of different classes (C1~C12) were crosswise calculated using the formula (1). Class i and class j were mutually contained and must be combined into the same class if $s_{ij} < 0$. Then, the separation degree g_{ij} between the sample sets of different classes which have been combined was crosswise calculated by the formula (2), and the separation degree matrix M_0 was generated by the g_{ij} above. All classes in M_0 were be used as the leaf nodes of the binary decision tree.

The minimum value (non-negative) of the separation degree was sought from the values of non-diagonal location in M_0 . The appropriate kernel function (RBF) and related parameters were chosen, and the training sample sets of two classes (C_m and C_n) corresponding to the minimum value were trained. Then the first internal node SMO1 (two classes' SMO classifiers) was generated and was saved into the list of internal nodes of the binary decision tree. Class C_m and class C_n were combined into C_{mn} , and the separation degree between C_{mn} and other classes were calculated again using the formula (2). At last, the separation degree matrix was updated as M_1 .

The minimum value (non-negative) of the separation degree was also sought from the values of non-diagonal location in M1. The appropriate kernel function (RBF) and related parameters were chosen, and then the training sample sets of two classes (C'_m and C'_n) corresponding to the minimum value were trained. Then the second internal node SMO2 (two classes' SMO classifiers) was generated and added into the list of internal nodes of the binary decision tree. Class C'_m and class C'_n were combined into C'_{mn} , and the separation degree between C'_{mn} and the other classes was calculated again using the formula (2). At last, the separation degree matrix was updated as M2.

The above steps were repeated until all classes were combined as the same one, and all internal nodes were generated. The classification model of multi-class SMO based on the binary decision tree was ultimately generated. The BDT-SMO training process based on the separation degree is as shown in Fig 5.

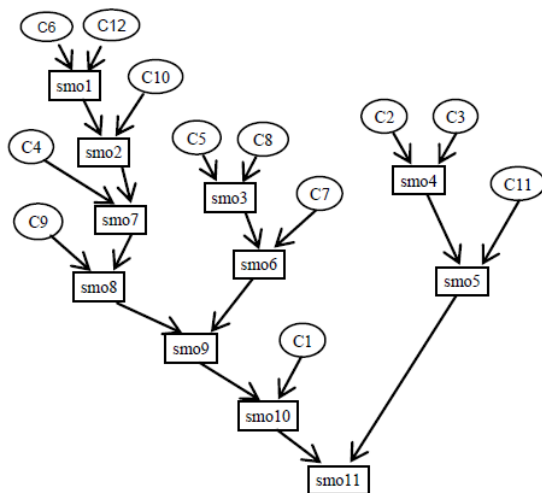


Figure 5. The BDT-SMO training process based on separation degree

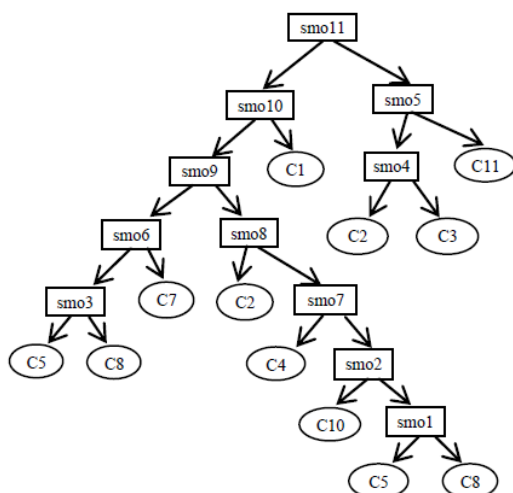


Figure 6. BDT-SMO classification test process based on separation degree

BDT-SMO classification test process based on the separation degree is opposite to the above training process. Test samples were processed starting from the root node, and each sample was judged by internal nodes

whether it belonged to the left or right branch. The class of the leaf node eventually reached was the class which the test sample belonged to. The specific classification test process is shown in the Fig 6.

Besides BDT-SVM, the traditional SVM can also be replaced with SMO with other multi-class SVM classification methods. a variety of different multi-class area based on combined features, and the result is shown in table II. The experimental result shows that the classification accuracy of BDT-SMO classifier based on separation degree is higher than that of other four methods; the test time and total time (including training time and testing time) of BDT-SMO classifier is better than other four methods although the training time of BDT-SMO classifier is not the best.

III. EXPERIMENTAL RESULT AND ANALYSIS

Select two different sets of samples which are based on spectral features and spectrum-texture combined features, and use the four different methods of maximum likelihood, BP neural network, BDT-SVM and BDT-SMO to train the two different sets of samples above separately, so as to verify the superiority of the BDT-SMO classification method based on the combined features.

A. BDT-SMO Classification Results and Accuracy Analysis Based on Combined Features

The BDT-SMO classifier based on separation degree was trained and applied to the land-use status classification of hyperspectral images in experimental area. The classification results are shown in Fig 7.

SMO classification methods were applied to the classification of the land-use status in the experimental area. The accuracy of classification results was verified by test samples and counted by confusion matrix, the classification accuracy of different types of ground objects can be analyzed with user accuracy (UA) and producers accuracy (PA), meanwhile the overall classification accuracy of the experimental area can be analyzed with overall accuracy (OAA) and Kappa coefficient. In addition, the classification efficiency of different methods can be analyzed with the test time. The analysis results are shown in table III and table IV.

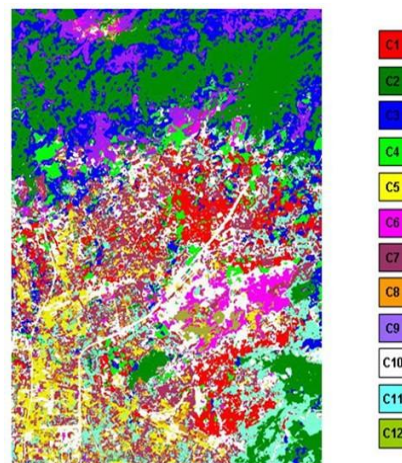


Figure 7. BDT-SMO classification results based on combined features

TABLE II. THE COMPARISON ON THE PERFORMANCES OF DIFFERENT MULTI-CLASS SMO CLASSIFICATION METHODS

Classification method	Overall accuracy (%)	Kappa coefficient	Training time(s)	Test time(s)
OAO-SMO	80.61	0.8446	24.54	21.20
OAA-SMO	79.02	0.8031	25.33	23.93
DAG-SMO	79.17	0.8210	30.59	15.47
BDT-SMO	81.61	0.8676	30.82	13.55
ECC-SMO	80.17	0.8395	31.35	14.83

TABLE III. THE EVALUATION ON CONFUSION MATRIX AND CLASSIFICATION RESULT

Land-use Type	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	UA(%)
C1	230	0	0	0	0	0	0	0	0	10	8	6	90.55
C2	0	231	0	0	0	0	0	0	0	0	0	0	100.00
C3	0	50	200	0	0	0	0	0	0	0	13	0	76.05
C4	2	0	0	222	0	7	0	0	0	19	1	0	88.45
C5	0	0	0	0	161	0	1	40	0	45	19	0	60.53
C6	0	5	0	0	6	222	0	5	0	25	0	20	78.45
C7	0	6	0	0	6	0	236	9	0	15	6	0	84.89
C8	0	0	0	0	30	0	9	189	0	15	9	0	75.00
C9	0	0	0	0	0	0	0	22	231	8	0	0	88.51
C10	0	6	0	0	20	0	0	12	0	227	7	0	83.46
C11	5	14	0	0	0	0	0	0	0	5	229	0	90.51
C12	0	0	0	0	5	35	0	5	0	42	0	165	65.48
PA(%)	97.05	74.04	100.00	100.00	70.61	84.09	95.93	67.02	100.00	55.23	78.42	86.39	OAA =81.61% Kappa =0.7994

TABLE IV. COMPARATIVE EXPERIMENTAL ANALYSIS RESULT

Classification Method	Overall accuracy (combined textures)	Kappa (combined textures)	Test time (combined textures)	Overall accuracy (spectrum)	Kappa (spectrum)	Test time (spectrum)
MLD	73.21%	0.7235	21.45s	73.46%	0.7004	14.78s
BPN	78.42%	0.7431	20.33s	75.87%	0.7378	13.65s
BDT-SVM	81.12%	0.7891	16.44s	77.14 %	0.7475	11.85s
BDT-SMO	81.61%	0.7994	13.55s	78.11%	0.7575	10.34s

The above classification result shows that the classification effect of BDT-SMO based on combined features was good in the experimental area, the overall classification accuracy was up to 81.61%, and Kappa coefficient was 0.7994. According to the user accuracy (UA) and producer accuracy (PA) of each class of land-use status, the classification accuracy of garden land (C3) was low (76.05%), and part of the gardens was classified into the woodland (C2) by mistake. That was because that garden land and woodland had no obvious boundaries, and there are lots of bushes in the border zone between them. Therefore, it was more difficult to distinguish the two class of land-use status accurately; the classification accuracies of the commercial take land (C5), public administration and public service land (C8), and other lands (C12) were lower and the specific values were 60.53%, 75.00%, and 65.48% respectively. That was mainly because the national latest land-use status classification standard (GB/T 21010-2007) was not all based on the geographical spatial features of ground objects. The classes of C5, C8, and C12 were classified mainly according to the land-use functions, their surface composition was complicated, and their geographical spatial features were not obvious. Therefore, it was difficult to achieve the ideal effect if the land-use status was classified simply using remote sensing images.

B. Comparative Analysis

Select two different sets of samples which are based on spectral features and spectrum-texture combined features,

and use the four different methods of maximum likelihood, BP neural network, BDT-SVM and BDT-SMO to train the two different sets of samples above separately. The classification effects of the different methods were analyzed with overall accuracy, Kappa coefficient, and test time. The experimental results are shown in table IV.

From table IV, it was known that the classification methods based on combined features consumed more time, but all classification methods based on combined features have higher classification accuracy than that of using simple spectral features except the maximum likelihood method (MLD). The classification accuracies of BP neural network (BPN), BDT-SVM, and BDT-SMO were increased by 2.55%, 3.98%, and 3.98% respectively, while the excessive learning problem appeared in the maximum likelihood method (MLD) based on combined features so that the classification accuracy was reduced by 0.25%. Under the two different condition of using combined features and spectral features, the classification effects of BDT-SVM and BDT-SMO are all significantly better than that of BPN and MLD; the classification accuracy and speed of BDT-SMO are all significantly higher than that of BDT-SVM; the classification accuracy and speed of MLD are the lowest. Therefore, the best effect can be obtained if the land-use status of the experimental area is classified using hyperspectral remote sensing images based on combined features.

C. Analysis and Discussion

(1) In the training of a multi-class SMO classifier, the best penalty factor C and related kernel parameters are determined by cross validation, and more time will be spent under the condition of large samples and dimensions. This is one of the important factors affecting the SMO classification efficiency. In the study of the future, the above parameters will be determined by better global optimization methods such as genetic algorithm.

(2) In the BDT-SMO classification method based on combined features, only the spectral features, textural gray level and gradient distribution features, and NDVI are combined, while other features such as spatial scale, elevation and topography information of the ground objects are not considered. Therefore, the classification accuracy will be further improved by more features.

IV. CONCLUSION

In this paper, the land-use status of the experimental area was classified comprehensively using spectral features, textural features, and BDT-SMO algorithm based on separation degree. It was a better solution to the problems of hyperspectral remote sensing imagery such as many bands, larger amount of data, high proportion of mixed pixels and lower spatial resolution and so on, and the classification effect was good. Select two different sets of samples which are based on spectral features and spectrum-texture combined features, and use the four different methods of maximum likelihood, BP neural network, BDT-SVM and BDT-SMO to train the two different sets of samples above separately. The experimental results showed that the overall accuracy of the BDT-SMO land-use status classification of experimental area based on combined features was up to 81.61%, and Kappa coefficient reached 0.7994. Thus, its classification accuracy and speed are significantly higher than those of other methods, and the land-use status classification efficiency of the experimental area can be effectively improved. For this reason, BDT-SMO classification method is with good generalization ability. Certainly, the BDT-SMO classification method based on combined features has some disadvantages currently, and therefore related problems should be further discussed in the future studies.

ACKNOWLEDGMENT

The work was supported by the National natural science foundation of China (NO. 41201417) and the Science and Technology Project of Fujian Province

Education Department of China (NO. JA13364). The author would like to thank Zhilei Lin and Professor Luming Yan for useful assistances, suggestions and discussions.

REFERENCES

- [1] D. F. Wang, Z. G. Yang, A. S. Wei. "Application of Texture Information on the Classification of Remote Sensing Imagery", *Journal of Nanjing Forestry University (Natural Science Edition)*, vol. 3, pp. 97-100, 2010.
- [2] J. liang, Z. K. Zhang, J. Li, et al. "Research on the Extraction of Vegetation Based on Gray Level-grads Co-occurrence Matrix", *Hydrographic Surveying and Charting*, vol. 1, pp. 29-31, 2013.
- [3] H. Zhang, J. L. Fan, et al. "Maximum Weighted Conditional Entropy Threshold Algorithm based on Gray-gradient Co-occurrence Matrix Model". *Computer Engineering and Application*, vol. 6, pp. 10-13, 2010.
- [4] S. G. Luo, C. R. Luo, Y. M. Zhou. "Training Strategies for Large-scale Sample Sets", *Journal of Guangdong Polytechnical Normal University*, vol. 9, pp. 30-33, 2008.
- [5] X. G. Zhang. "About the Statistical Learning Theory and Support Vector Machine (SVM)", *ACTA AUTOMATICA SINICA*, vol. 1, pp. 37-38, 2000.
- [6] M. K. Li, J. L. Wang, M. X. Wang, et al. "Mixed Pixel Classification of Multi-spectral Image Based on Binary Decision Tree". *Geospatial Information*, vol. 3, pp. 100-102, 2010.
- [7] G. Dai, W. Cui, Y. Zhang, et al. "Clustering Analysis of Weld Defect based on Gray-gradient Co-occurrence Matrix". *China Safety Science Journal (CSSJ)*, vol. 5, pp. 79-85, 2013.
- [8] S. Li, Q. A. Ren, B. Wen, et al. "Multi-class SVM with Binary Decision Tree based on Degree of Separation Applied in Welding Defect Classification and Recognition". *Journal of Sichuan University (Natural Science Edition)*, vol. 3, pp. 520-524, 2010.

Fenghua Huang, born in 1982. He is currently a lecturer in Sunshine College of Fuzhou University in China, and meanwhile a Ph.D. student in Geographical Sciences College of Fujian Normal University in China. He had received his Master degree in software engineering from Fuzhou University in 2009. His main research interests are high performance data mining, remote sensing and GIS applications.

Luming Yan, born in 1951. He is currently a professor and PhD Candidates Supervisor in Geographical Sciences College of Fujian Normal University in China. His main research interests are natural geography, remote sensing and GIS applications.