Abstract—A method combined with support vector machines (SVMs) and pairwise coupling (PWC) was developed to achieve land use/land cover fractions of a moderate-resolution remote sensing image. At first, SVMs were applied to solve classification problems. Then, they were extended with PWC to output probabilities as the abundance of landscape fractions. The performances were evaluated by using the “estimated” landscape class fractions from our method, fully constrained least squares method, and unmixing nonlinear SVM (u_NLSVM) method, respectively, and the results were validated by real fractions generated from the SPOT High Resolution Geometric (HRG) image. The best classification results were obtained by the proposed method, which proved the effectiveness of our method.

Index Terms—Pairwise coupling (PWC), pixel unmixing, support vector machines (SVMs).

I. INTRODUCTION

LAND COVER composition is an important factor for earth and environmental sciences. As most general land cover classes (e.g., forest, grassland, etc.) differ in their interaction with the radiations commonly used in remote sensing, a remote sensing image is well suited as a source of information on environmental features such as land cover [1]. Moderate-resolution satellite images such as Landsat Thematic Mapper (TM) and SPOT High Resolution Geometric (HRG) provide a distinctive opportunity to capture a landscape over a large area in a timely and cost-effective manner [2], [3]. Most traditional techniques are inherently full-pixel-based techniques. One basic assumption for these techniques is that each pixel vector measures the response of one predominant ground object. Unfortunately, in moderate-resolution satellite images, many ground sites corresponding to a pixel (e.g., 30 m × 30 m for a TM image) are mixtures of objects. It is unsurprising that land cover maps from pixel-based methods often contain considerable errors. To accurately map land cover, the interpretation of mixed pixels becomes a key factor in the analysis of remote sensing image [4].

Spectral mixture analysis (SMA) requires a previous step where the spectral signatures of ground constituents (endmembers) are identified by an supervised or unsupervised method, and is then used to estimate the abundance fractions of these signatures by expressing individual pixels as a linear or nonlinear combination of endmembers [5]–[7]. Several algorithms have been proposed over the last decade to find endmembers directly from the image data [4], [8]–[12]. The automated morphological endmember extraction (AMEE) method which integrates the spatial and spectral information is able to provide a relatively good characterization of general landscape conditions [4].

For spectral unmixing, a number of approaches have been developed to tackle the decomposition of mixed pixels such as linear SMA (LSMA) [13]–[16], extended support vector machine (SVM) [17], [18], fuzzy spectral unmixing [19], neural network mixture model [20]–[22], Bayesian spectral unmixing [3], [23], as well as gradient-based optimization algorithm [24]. The LSMA has so far been the most popular approach, given its simple mathematical form. For the LSMA-based method, in order to provide accurate and reliable estimates of signature abundance fractions for material quantification, sum-to-one and nonnegativity constraints must be imposed on the abundance fractions of materials in a pixel. The fully constrained least squares (FCLS) unmixing algorithm [25]–[27] is widely adopted as a practical solution.

SVMs have been recently introduced in the statistical learning theory domain [28], [29] for regression and classification problems and applied to the classification of multi- and hyperspectral images [30]–[33]. In this paper, we develop an SVM–PWC method to improve the estimating precision of land use/land cover fractions from a Landsat ETM+ image. Originally, the SVM-combined-with-PWC method is used to construct the multiclass classifier. Kolaczyk [34] has discussed the reason of Bayesian posterior probability instead of fractional abundances of endmember and concluded that subpixel classification (spectral unmixing) is a special case of
classification by posterior probability. PWC can provide a multiclass probability estimate by combining all pairwise comparisons. Kernel SVM can project the nonlinear data in a higher dimensional space where they are linearly separable to solve mixed pixels in nonlinear mixture condition, which are more often than in linear mixture condition. Therefore, the SVM-PWC method can be adopted to generate fractions of landscape physical constituents according to spectral variability. In order to examine the effectiveness of this method, we finally compare the classification results with FCLS and unmixing nonlinear SVM (u_NLSVM) methods.

II. STUDY SITE AND DATA PREPARATION

The study site is in the south of Guangdong, China (21°48′–22°27′ N, 113°3′–114°19′ E), enjoying a subtropical oceanic climate. The average annual temperature is around 22.1°C-23.4°C, with an average annual rainfall of 1770-2300 mm. This area has experienced great changes in the past 20 years and has been undergoing a quick urbanization process since the 1980s [35]. Based on the field experience and spectral attribute in the image, five typical classes can be recognized, including 1) water, 2) woodland, 3) shade, 4) residential area, and 5) industrial land. Shade, caused by tall buildings or trees, is an inevitable part of the urban area, particularly in the developed areas [2]. In this paper, we classified it as a single class to reduce the errors in the process of classification.

One scene of the Landsat ETM+ image (179 pixels × 111 pixels) acquired over Guangdong on November 20, 2004, was used as experimental data in this paper. To reduce atmospheric distortion, the digital number value at each band was first converted to the normalized at-sensor reflectance. Minimum noise fraction transformation was then applied to remove the correlation existing between the six bands (band 6, a thermal infrared band with different spatial resolution, was excluded). A 5-m-resolution image (1074 pixels × 666 pixels) from the SPOT HRG image released in March 2005 was adopted as a reference image in this paper. The Landsat ETM+ image released in March 2005 was adopted as a reference image in this paper. The Landsat ETM+ image was further registered to the SPOT HRG image, so the two images covered the same area. Fig. 1 shows the data of the ETM+ image and the corresponding SPOT HRG image.

III. METHODOLOGY

A systematic framework of our method is shown in Fig. 2. Three sets of proportion results were derived from experimental data using SVM-PWC, FCLS, and u_NLSVM, respectively. The same set of endmembers selected by AMEE was used in the SVM-PWC and u_NLSVM methods. FCLS can only use one endmember for each class, so the final set of endmembers for the FCLS method was generated by averaging endmembers selected by AMEE. The major difference between these methods lies in the way mixed pixels are decomposed. The classification result of the SPOT HRG image was adopted as reference for choosing the endmembers and assessing the classification accuracy of the proportion results from the SVM-PWC model. The unmixing accuracy was assessed among the SVM-PWC, FCLS and u_NLSVM results from the ETM+ image. The following sections will describe the methods of SVM and PWC in detail.

A. Spectral Unmixing Based on SVM

1) SVM: SVM classification algorithms, originally proposed by Vapnik [36] to solve two-class problems, are based on finding a separation between hyperplanes defined by classes of data. The complete mathematical formulation of SVM can be found in [37].

Fig. 5 shows fraction maps unmixed by the u_NLSVM method. The figure shows a better visual effect than Fig. 4 generated by FCLS for nonlinear SVM is valuable for unmixing data which have complicated class boundaries in the original feature space.

SVM can be generalized to compute nonlinear decision surfaces. The method consists in projecting the data in a higher dimensional space where they are linearly separable. In fact, the projection can be simulated using a kernel method. Here, the C-SVC method was used to solve the problem

$$\min_{w, b, \xi} \frac{1}{2} w^T w + C \sum_{i=1}^{l} \xi_i$$

subject to

$$y_i (w^T \Phi(x_i) + b) \geq 1 - \xi_i, \quad \xi_i \geq 0, \quad i = 1, \ldots, l.$$  (1)

Here, the training vector \(x_i\) is mapped into a higher (maybe infinite) dimensional space by the function \(\Phi\). Normally, the \(\Phi\)
function can be a linear function, a polynomial function, a radial
basis function, a sigmoid function, or a precomputed kernel
function. When extending SVM to solve multiclass problems,
the following two approaches can be used [38]: 1) one against
all: $k$ SVMs are iteratively applied on each class against all
the others, where $k$ is the number of classes [37], and 2) one
against one: $k(k-1)/2$ SVMs are constructed, where each one
is trained on data from two classes [39]. The second approach
is used in this paper.

2) Abundance Estimation: The SVM method can be com-
bined with PWC to provide fractional abundances of materials.
Several authors have proposed probability estimates by com-
bining pairwise class probabilities [40]–[42]. Here, we choose
the method proposed by Kolaczyk [42], which is more highly
robust to sample probability distribution and classification num-
ber than the pairwise coupling (PWC) method that Wu
et al. [17] used [42].

Given $k$ classes of data, for any $x$, the goal is to estimate
posterior probability

$$ p_i = p(y = i|x), \quad i = 1, \ldots, k. \quad (2) $$

Following the setting of the one-against-one approach for
multiclass classification, first, pairwise class probabilities are
estimated as follows:

$$ r_{ij} \approx p(y = i|y = i \text{ or } j, x) \quad (3) $$

using an improved implementation proposed by Platt [43]

$$ r_{ij} \approx \frac{1}{1 + e^{-A f + B}} \quad (4) $$

where $A$ and $B$ are estimated by minimizing the negative
log-likelihood function using known training data and their
decision values $f$. Labels and decision values are required to
be independent, so here, we conduct fivefold cross-validation
to obtain decision values.

Then, the PWC method by Wu et al. [42] is used to obtain $p_i$
from all $r_{ij}$. It solves the following optimization problem:

$$ \min_p \frac{1}{2} \sum_{i=1}^{k} \sum_{j \neq i} (r_{ji}p_i - r_{ij}p_j)^2 $$

subject to

$$ \sum_{i=1}^{k} p_i = 1, \quad p_i \geq 0 \quad \forall i. \quad (5) $$

Using (5), the abundance fractions generated by the SVM-
PWC method can satisfy sum-to-one and nonnegativity con-
straints. In addition, we choose the LibSVM simple iterative
method to simplify the implementation [44].

Before spectral unmixing, the attribute data of samples se-
lected by AMEE were scaled so as to fall within a small speci-
cified range, i.e., between 0 and 1. Normalizing the input values
for each attribute measured in the training samples can help not
only speed up the learning phase but also prevent attributes
with initially large ranges from outweighing attributes with
initially smaller range. Here, radial basis function was chosen
as kernel function for its classification recognition power is not
less than that of polynomial and sigmoid functions. Probability
parameter $b$ was set at 1 for getting abundance. After cross-
validation computing, we got the results of $C = 2048$ and
$\gamma = 0.0078125$.

B. Accuracy Assessment

Accuracy assessment is an essential step for landscape map-
ping from remotely sensed data [45]. All pixels from the SPOT
HRG image were chosen as an independent set of test samples.
For each test sample unit, its actual fraction of each land use
type was acquired through the SPOT HRG classification image
(Fig. 3).

A conventional confusion matrix is inappropriate for sub-
pixel classification evaluation since multiple memberships exist
with each pixel. In this paper, we adopted two methods to
evaluate the accuracy of land cover/land use composition for

![Fig. 3. SPOT HRG classification image generated by the MLC method.](image-url)
Fig. 4. Fractional images generated by the FCLS method.

Fig. 5. Fractional images generated by the u_NLSVM method.

Fig. 6. Fractional images generated by the SVM–PWC method.

<table>
<thead>
<tr>
<th>TABLE I</th>
<th>COMPARISON OF MAE VALUES BETWEEN SVM–PWC, FCLS, AND u_NLSVM</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAE</td>
<td>woodland</td>
</tr>
<tr>
<td>SVM-PWC</td>
<td>0.19</td>
</tr>
<tr>
<td>FCLS</td>
<td>0.41</td>
</tr>
<tr>
<td>u_NLSVM</td>
<td>0.22</td>
</tr>
</tbody>
</table>
TABLE II
CROSS TABLE OF SVM-PWC, FCLS, AND u_NLSVM

<table>
<thead>
<tr>
<th>reference data</th>
<th>woodland</th>
<th>water</th>
<th>shade</th>
<th>residential area</th>
<th>industry land</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVM-PCW</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>woodland</td>
<td>33.12</td>
<td>0.76</td>
<td>0</td>
<td>0</td>
<td>0.87</td>
</tr>
<tr>
<td>water</td>
<td>0</td>
<td>12.22</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>shade</td>
<td>0</td>
<td>0.17</td>
<td>4.65</td>
<td>0</td>
<td>0.19</td>
</tr>
<tr>
<td>residential area</td>
<td>0</td>
<td>2.31</td>
<td>0</td>
<td>34.34</td>
<td>2.63</td>
</tr>
<tr>
<td>industry land</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>8.74</td>
</tr>
<tr>
<td>Overall accuracy</td>
<td>0.93</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FCLS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>woodland</td>
<td>21.28</td>
<td>1.70</td>
<td>6.40</td>
<td>0</td>
<td>3.74</td>
</tr>
<tr>
<td>water</td>
<td>0</td>
<td>15.45</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>shade</td>
<td>0</td>
<td>0</td>
<td>4.65</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>residential area</td>
<td>0</td>
<td>2.44</td>
<td>9.18</td>
<td>17.37</td>
<td>5.36</td>
</tr>
<tr>
<td>industry land</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>12.44</td>
</tr>
<tr>
<td>Overall accuracy</td>
<td>0.71</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>u_NLSVM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>woodland</td>
<td>33.12</td>
<td>0.29</td>
<td>0.13</td>
<td>0</td>
<td>1.39</td>
</tr>
<tr>
<td>water</td>
<td>0</td>
<td>13.97</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>shade</td>
<td>0</td>
<td>0</td>
<td>3.96</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>residential area</td>
<td>0</td>
<td>1.19</td>
<td>0.55</td>
<td>34.34</td>
<td>5.74</td>
</tr>
<tr>
<td>industry land</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5.298</td>
</tr>
<tr>
<td>Overall accuracy</td>
<td>0.91</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Each land class [2]. The first method is the mean absolute error (MAE) [46], [47], which is defined as follows:

\[
MAE_i(1^{P_i}, 2^{P_i}) = \frac{1}{N} \sum_{j=1}^{N} |1^{P_i,j} - 2^{P_i,j}| \tag{6}
\]

\[
Total_{MAE} = \frac{1}{2} \sum_{i=1}^{J} MAE_i \tag{7}
\]

where \(1^{P_i,j}\) is the estimated land use composition fraction from SVM-PWC classification, \(2^{P_i,j}\) is the "actual" land use composition fraction digitized from SPOT HRG classification for \(i\) in the test sample \(j\), \(N\) is the total number of test samples, and \(J\) represents the number of categories.

The generalized cross-tabulation matrix, as introduced by Pontius and Cheuk [48], was also adopted to further analyze the disagreement between the estimated results and reference maps. Specifically, a composite operator was used to calculate the entries in the cross-tabulation matrix through (8), shown at the bottom of the page, where \(P_{i,j}\) is the estimated entry in the cross-tabulation matrix for classes \(i\) and \(j\). \(1^{P_i}\) represents the percentage of the class \(i\) from SVM-PWC classification, and \(2^{P_i}\) is the proportion of the landscape for which the class \(i\) exists at the reference SPOT HRG classification map. Once all \(P_{i,j}\) entries were filled, overall accuracy and Kappa values were calculated in the same manner as that applied in traditional classification.

IV. RESULTS AND ANALYSIS

Fig. 3 shows the classification result of the SPOT HRG image, which was classified with the maximum likelihood classification (MLC) method. The accuracy of each classification was checked with the random sampling method, and 25 samples were selected for each land cover.

\[
P_{i,j} = \begin{cases} 
\min \left(1^{P_i}, 2^{P_i}\right), & \text{if } i = j \\
\left[1^{P_i} - \min \left(1^{P_i}, 2^{P_i}\right)\right] \times \left[\frac{2^{P_i} - \min \left(1^{P_i}, 2^{P_i}\right)}{\sum_{j=1}^{J} (2^{P_i} - \min \left(1^{P_i}, 2^{P_i}\right))}\right], & \text{if } i \neq j 
\end{cases} \tag{8}
\]
category to check the ground truth. The sample data were collected from the published land cover maps. The average accuracy of classification and the Kappa coefficients were determined to be 0.81 and 0.78, respectively. Without higher accuracy classification map and ground truth data, the classification result of the SPOT HRG image was adopted as reference for choosing the endmembers as well as the test samples to assess the unmixing accuracy of the proportion map from the SVM-PWC method.

Fig. 4 shows five fraction images unmixed from the ETM+ image using the FCLS method. The gray level was used to designate the percentage of a specific class that a pixel contains. Therefore, a brighter pixel means a higher fraction. There are obvious misinterpretations in shade and residential area. The image fraction maps estimated by the proposed algorithm of this paper are shown in Fig. 6. These figures also show a better visual effect than that generated by FCLS. Particularly, fractions of the shade and residential area have been refined. Moreover, the bright pixels in the industrial, water, and woodland fraction images correspond well to the original ETM+ image. The fact that SVM-PWC performs better than the FCLS method can be attributed to two reasons. First, SVM-PWC is able to separate classes which are very close to each other with a number of training samples, while the FCLS method can only use one endmember for each class to analyze a mixture pixel. In a moderate-resolution satellite image, one land cover class may include more than one spectrum because of slope, direction, land cover type, or other reasons. Thus, the multiple endmembers used to decompose a mixture pixel can get better results. Second, SVM-PWC can resolve a nonlinear mixture pixel using a kernel method to simulate the nonlinear projection of data in a higher dimensional space. Due to the reason that SVM decomposes a nonlinear mixture pixel and gets positive unmixing results without any artificial constraints, the visual effects of both SVM-PWC and u_NLSVM are all excellent.

Table I shows the accuracy indices (MAE) of SVM-PWC, FCLS, and u_NLSVM. Here, a smaller index denotes a better unmixing result. We can easily find that SVM-PWC performed better than the other two methods since it has the smallest total MAE (0.33) in Table I. The best fraction is associated with the water category, showing the smallest MAE (0.06); the second best fraction is the industrial category; and the worst fraction is the residential class. A possible explanation could be that residential and woodland, due to the spectral confusion of various surface types, generated worse results. Reference data may bring some misjudgments because of their general incorrectness. Nevertheless, we can still find that the SVM-PWC method can get better results under the same conditions.

Table II shows the cross-tabulation matrix for SVM-PWC, FCLS, and u_NLSVM. Evidently, SVM-PWC showed an
obvious improvement in the results with the highest overall accuracy (0.93) and Kappa (0.90).

Figs. 7–9 present the plots showing the relationships for 30 test samples between the estimated fraction using three methods and the “actual” fraction digitized from the reference SPOT HRG image. In this paper, we randomly selected 30 test samples from the ETM+ image. For each sample, the actual fraction of each landscape class was acquired from the SPOT HRG classification image. The distance that the test samples deviate from the line $y = x$ was then used to assess the closeness between the actual and estimated proportions of each class [2]. The total distance (total D) of each class is also shown in each plot. In these plots, the closer the test samples to the line are, the better the estimated result will be.

The correlation between the estimated fraction using FCLS and the actual fraction from the SPOT HRG image is shown in Fig. 7. In contrast to the SVM-PWC method, a lower correlation can be discerned, although most observed points distribute along the $y = x$ line.

Compared to FCLS, SVM-PWC has achieved a promising result with the smaller overall distance to the $y = x$ line (Fig. 8). As previously discussed, such a situation may be a result of the ability from PWC schemes to make nonlinear spectral mixture of pixels and get posterior probability. Comparing the SVM-PWC with the u_NLSVM (Fig. 9) method, the total distance got by the u_NLSVM is higher than that by the SVM-PWC method in woodland and residential area, but lower than that by the SVM-PWC method in water, shade, and industry land. The differences of total distance got by those two methods in water and shade are only 0.49% and 0.25%, respectively. Excluding the impact of sample selection, the SVM-PWC and u_NLSVM methods give the same accurate estimation of the landscape composition in the study area.

V. Conclusion

Using moderate-resolution satellite images, traditional pixel-based classification methods are usually inappropriate for landscape analysis since one pixel is always a mixture of multiple objects. In this paper, we have developed the SVM-PWC method to estimate the subpixel fractions for landscape mapping, and the performance of our method was compared to FCLS and u_NLSVM methods. The endmembers used in these methods based on ETM+ data were generated by the AMEE method. Our results show that the multiple endmembers used to estimate the abundance fractions can give a finer characterization of the landscape since it usually has various land covers of one class. Previous studies proved that the FCLS method was widely used in subpixel classification. In particular, it could effectively detect and quantify materials in hyperspectral
imagery. However, the straightforward application of FCLS was unsatisfactory to make nonlinear spectral mixture of pixels. SVM is also extended to map mixed pixels using hyperspectral data such as u_NLSVM. Although in some experiments, the SVM-PWC method performs the same accuracy estimation with the u_NLSVM method, we propose a new solution to derive the final fractions in this paper. Experiments turn out that the SVM-PWC method can provide consistently better results.

In this paper, although the SVM-PWC method achieved considerable improvements in landscape classification, unmixing results still have a misclassification problem. To get high accuracy, we could incorporate other information of an image, such as slope and texture, into our method. Moreover, how to set optimal parameters to SVM-PWC is another important issue for successful application of unmixing. Future work will investigate how to combine all those ideas in an SVM-PWC classification to develop a more sophisticated soft method.

ACKNOWLEDGMENT

The authors would like to thank Prof. B. de Jong of Utrecht University for his help to improve the English of this paper. The authors would also like to thank the three anonymous reviewers and the Associated Editor of this journal for their constructive comments, which have helped improve greatly the readability of this paper. This is contribution No. IS-1370 from GIGCAS.

REFERENCES

Abstract

In this paper, we present a new approach to analyze the spatial restructuring of land use patterns. We apply a statistical learning model to extract meaningful patterns from remote sensing data. The method involves training a support vector machine (SVM) classifier on a small set of labeled training data. The classifier is then used to predict land use classes in an unsupervised manner, using a regularization approach to estimate the posterior probabilities of the classes.

The results show a significant improvement in the classification accuracy compared to traditional linear unmixing techniques. The method is particularly useful in areas with complex land use patterns and limited training data.

Keywords: statistical learning, support vector machine, land use classification, remote sensing.
Yan Li was born in Xinjiang, China, in 1955. She received the B.S. degree in geography from Xinjiang University, Urumqi, China, in 1977 and the M.S. degree in applied geography information system from the Faculty of Geo-Information Science and Earth Observation (ITC), Enschede, The Netherlands, in 1992. In 1977, she was with the Remote Sensing Section, Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, Urumqi, and transferred to the Remote Sensing and GIS Center, Guangzhou Institute of Geography, Guangzhou, China, and the Guangdong Academy of Sciences, where she has been a Research Professor with the institute since 2000. She is currently a Professor of computer and the Director of the Spatial Information Research Center, South China Normal University, Guangzhou. She was also a Visiting Scholar at the University of California, Santa Barbara, and Slippery Rock University, Slippery Rock, PA. Her main research interests include remote sensing and GIS applications on environment and natural resources.

Prof. Li was a recipient of the Second-Rank Award of the National Sciences and Technology Progress of China and the First-Rank Award of Sciences and Technology of Guangdong Province.

Xingfang Wang was born in Shanxi, China, in 1971. She received the B.S degree in computer science from the Beijing University of Technology, Beijing, China, in 1994, the M.S degree in computer science from Jinan University, Guangzhou, China, in 2004, and the Ph.D. degree in environmental science from the Guangzhou Institute of Geochemistry, Chinese Academy of Sciences, Guangzhou, in 2011. She has been with South China Normal University, Guangzhou, since 2002, where she has been a Lecturer since 2004. Her research interests include remote sensing and applications on environment.