2014 IEEE International Conference on Systems, Man, and Cybernetics
October 5-8, 2014, San Diego, CA, USA

Multi-scale Patch based Box Kernels for Hyperspectral Image Classification

Jiangtao Peng
Faculty of Mathematics and Statistics,
Hubei University
Wuhan, China
Email: pengjt1982@126.com

Yicong Zhou, Member, IEEE
Department of computer and Information Science,
University of Macau
Macau, China
Email: yicongzhou@umac.mo

C.L. Philip Chen, Fellow, IEEE
Department of computer and Information Science,
University of Macau
Macau, China
Email: philipchen@umac.mo

***Abstract*—Integrating labeled pixels with prior knowledge of hyperspectral spatial homogeneous regions, we propose a region-based hyperspectral image classification method, called the support vector machine with the multi-scale patch based box kernel (SVM-MPBK). It models the local homogeneous region of each pixel as a box, and measures the similarity between different box regions using box kernel. The box is represented as multidimensional intervals computed band by band in a neighborhood pixel patch. Using multi-scale patches to calculate box, SVM-MPBK fuses the complementary classification results in different scales by a majority voting. Experimental results on benchmark hyperspectral data sets demonstrate the effectiveness of SVM-MPBK.**

***Index Terms*—Support Vector Machine, box kernels, hyperspectral image, classification, multi-scale**

I. Introduction

Hyperspectral image (HSI) data contain a set of images of the same geographical scene. These images correspond to different spectral bands of electromagnetic radiation. Fixed a band, it obtains a scene image containing spatial structure information. Fixed an image coordinate position, it obtains a spectral curve vector, also called a pixel. Different materials have different absorption and reflection at a certain spectral band. Thus, it can identify and classify the materials based on the spectral curves or pixels. Traditional classifiers, such as support vector machine (SVM) and other kernel-based methods, are all using the spectral signatures for HSI classification and have shown good performance [1], [2], [3].

In SVM, each labeled pixel is considered as a sample and processed independently. It does not exploit the correlations between spatial neighboring pixels. For HSI, neighboring pixels have similar spectral characteristics and usually belong to the same class of materials. The relevance between neighboring pixels can be used to improve the SVM classification performance [4], [5], [6], [7]. SVM with the composite kernel (SVM-CK) method incorporates spatial contextual information to design a new composite kernel, which is easy to implement and much effective [4]. In SVM-CK, each pixel is considered as not only spectral features but also spatial features. The spatial feature is represented as the mean or standard deviation of pixels in a spatial pixel neighborhood. Then, based on the obtained spectral and spatial features, the corresponding spectral and spatial kernels can be computed. Finally, a composite kernel is formed by a linear combination of these two kernels and can be used in conventional SVM methods. SVM-CK extracts spatial information of neighboring pixels via a mean (or standard deviation) vector. It is actually a sample-point-based classifier, in which a mean sample can not fully capture spatially pixel variations in a local patch.

Because an HSI contains many homogeneous regions, the pixels in a homogeneous region generally belong to the same class (see Fig. 5(a)). An intuitive idea is to classify these homogeneous regions directly rather than classifying the mean or standard deviation features extracting from the regions. In this paper, using the prior distribution characteristics of the HSI spatial regions, we propose a region-based HSI classification method, called the SVM with the multi-scale patch based box kernel (SVM-MPBK).

To our best knowledge, region based classification, using regions as features directly, has not been employed in the HSI data. The key issues of the region or set based SVM-MPBK classification are how to define a local homogeneous region and then how to measure the similarity between different regions. We model the local homogeneous region of each pixel as a box which consists of multidimensional intervals computed band by band in a neighborhood pixel patch. Then, the box kernel proposed in [8] is used to measure the similarity between different box regions. The box can capture spatial local pixel variations and the box kernel provides an accurate similarity metric by incorporating the spatial texture structure from the label boxes. Because the interval and box are sensitive to neighborhood patch sizes, we use multi-scale patches and fuse multi-scale classification results by a majority voting. The ensemble of the multi-scale complementary information ensures the excellent classification performance of SVM-MPBK.

II. Related work

The classic SVM methods deal with the labeled points, i.e., $\mathcal{L} = \{(\boldsymbol{x}_i, y_i), i = 1, \ldots, \ell\}$ with $\boldsymbol{x}_i \in \mathcal{X} \subset R^d$ and $y_i \in \mathcal{Y} = \{-1, 1\}$. The box kernel method deals with both the labeled points and labeled regions [8]. A region of the input space is

978-1-4799-3840-7/14/$31.00 ©2014 IEEE

represented by multidimensional intervals (box),

$$\boldsymbol{S}_j = \{\boldsymbol{x} \in R^d : x^z \in [a_j^z, b_j^z],\ z = 1, \dots, d\}.$$

The label of this region is $y_j \in \mathcal{Y}$.

The general kernel function measures the similarity between sample points. The box kernel extends the general kernel from points to sets. It can be computed by convolving the general kernel with the characteristic function of the box [8]. For Gaussian kernel $K_{pp}(\boldsymbol{x}, \boldsymbol{y}) = \exp(-\|\boldsymbol{x} - \boldsymbol{y}\|^2/(2\sigma^2))$, the set-to-point box kernel and set-to-set box kernel are computed as follows:

$$K_{sp}(\boldsymbol{S}_j, \boldsymbol{x}) = \frac{1}{\nu_j} \prod_{i=1}^{d} \frac{\sqrt{2\pi}\sigma}{2} \Big(h\Big(\frac{x^i - b_j^i}{\sqrt{2}\sigma}\Big) - h\Big(\frac{x^i - a_j^i}{\sqrt{2}\sigma}\Big) \Big)$$

$$\begin{aligned} K_{ss}(\boldsymbol{S}_p, \boldsymbol{S}_q) = \frac{(\sqrt{\pi}\sigma^2)^d}{\nu_p \nu_q} \prod_{i=1}^{d} \Big(& \phi(b_p^i, b_q^i) - \phi(a_p^i, b_q^i) \\ & - \phi(b_p^i, a_q^i) + \phi(a_p^i, a_q^i) \Big) \end{aligned}$$

where

$$\phi(a, b) := uh(u) - \frac{1}{\sqrt{\pi}} e^{-u^2}, \quad u = \frac{a - b}{\sqrt{2}\sigma},$$

and h is the complementary error function.

III. Proposed method

A. SVM-MPBK

SVM-MPBK aims to perform classification on the labeled pixels and labeled regions of HSI. One key issue is how to design the labeled box regions.

We first define a local pixel patch for a labeled pixel $\boldsymbol{x}_i$ as: $\boldsymbol{P}_i = \{\boldsymbol{x}_i, \boldsymbol{x}_{i1}, \boldsymbol{x}_{i2}, \dots, \boldsymbol{x}_{is}\}$, where $\boldsymbol{x}_{i1}, \dots, \boldsymbol{x}_{is}$ are the neighboring pixels of $\boldsymbol{x}_i$. Based on the patch $\boldsymbol{P}_i$, we determine multidimensional intervals to form a box region. The interval must faithfully reflect the spatial variation of homogeneous pixels. Thus, we eliminate unrelated background points and noisy pixels in $\boldsymbol{P}_i$ in advance. For this purpose, we compute the distance between the central pixel $\boldsymbol{x}_i$ and its neighboring pixels: $d_k = \|\boldsymbol{x}_i - \boldsymbol{x}_{ik}\|^2$, and delete 10% samples in $\boldsymbol{P}_i$ with the largest distances. Denote the reduced patch as $\hat{\boldsymbol{P}}_i = \{\hat{\boldsymbol{x}}_i, \hat{\boldsymbol{x}}_{i1}, \dots, \hat{\boldsymbol{x}}_{i\hat{s}}\}$.

Then, we set the multidimensional intervals band by band. Each band contains $\hat{s}+1$ numbers. A simple way to define an interval is using the minimum and maximum of these numbers. However, the extreme interval fails to reflect the variation of most samples and may enlarge the region. In order to obtain a mild interval, we set the lower and upper bounds of the interval as the 25-th and 75-th percentiles of the values in the $\hat{s}+1$ numbers, respectively. Considering all bands, we obtain a lower bound vector $\boldsymbol{a}_i$ and an upper bound vector $\boldsymbol{b}_i$ of the patch $\hat{\boldsymbol{P}}_i$. Finally, the box corresponding to the pixel $\boldsymbol{x}_i$ is

$$\boldsymbol{S}_i = \boldsymbol{a}_i \times \boldsymbol{b}_i = [a_i^1, b_i^1] \times [a_i^2, b_i^2] \times \cdots [a_i^d, b_i^d]$$

For each labeled pixel, we can exploit the spatial neighborhood information to build a labeled box set. Given a collection of training samples and corresponding labeled boxes

$$\mathcal{L} = \{(\boldsymbol{x}_1, y_1), \dots, (\boldsymbol{x}_\ell, y_\ell)\},$$
$$\boldsymbol{B} = \{(\boldsymbol{S}_1, y_1), \dots, (\boldsymbol{S}_\ell, y_\ell)\},$$

we can build a SVM classifier on these 2ℓ labeled pairs. In the training phase, we compute the kernel K,

$$K = \begin{bmatrix} K_{pp}(\mathcal{L}, \mathcal{L}) & K_{sp}^{\mathrm{T}}(\boldsymbol{B}, \mathcal{L}) \\ K_{sp}(\boldsymbol{B}, \mathcal{L}) & K_{ss}(\boldsymbol{B}, \boldsymbol{B}) \end{bmatrix}$$

and train SVM model to obtain the coefficient vector $\boldsymbol{\alpha}$ satisfying $\boldsymbol{f} = K\boldsymbol{\alpha}$. In the prediction phase, for a testing sample $\hat{\boldsymbol{x}}$, we compute the corresponding box $\hat{\boldsymbol{S}}$ and box kernel $\hat{K}$,

$$\hat{K} = \big[K_{sp}(\hat{\boldsymbol{S}}, \mathcal{L}) \quad K_{ss}(\hat{\boldsymbol{S}}, \boldsymbol{B}) \big],$$

and predict the label by using the decision value $\hat{f} = \hat{K}\boldsymbol{\alpha}$.

From the above procedures, we can see that the neighboring patch size (or the number of neighbors s) directly decides the box computation. It seriously affects the classification performance. The patch should cover HSI homogeneous regions as accurately as possible such that the computed intervals and boxes could faithfully reflect spatially local pixels variations. It is usually difficult to pre-define an optimal patch because an HSI usually contains small and large homogenous regions simultaneously and one single patch can not differentiate different pixel contextual structures. Moreover, for different HSI data sets, the optimal scale also varies a lot.

In order to alleviate the effect of the patch sizes, we compute the box kernel and perform the classification based on multi-scale patches. In detail, we first provide several patches and compute the patch-based box kernels, then perform SVM classification on each patch-based box kernel individually. Finally, the multi-scale classification results are fused by a majority voting. In the proposed multi-scale framework, it is free of the scale selection. Moreover, multi-scale patches can generate multi-scale boxes. The complementary information in different scales and boxes helps to accurately describe the local homogeneous regions and improve classification performance.

By incorporating the spectral labeled training points and spatial multi-scale labeled box sets, SVM-MPBK is summarized in Algorithm 1.

B. Discussion

In the proposed SVM-MPBK, we should note that:

1) Using a box to model the local homogeneous region is feasible. The box is used to reflect spatially local pixel variations. The only requirement is that the pixels in the box have the same label. Due to the region distribution characteristics of HSI, the local pixel patch or box usually contains the same material. Moreover, the combination of multi-scale boxes can approximate different irregular homogeneous regions.
2) Outlier elimination is necessary. The local patch of a pixel may contain noisy or background points, especially

Algorithm 1 SVM-MPBK

Input: Training samples $\{(\boldsymbol{x}_j, y_j)\}_{j=1}^{\ell}$, p patch scales
Training phase:
for $i = 1$ to p **do**
  for $j = 1$ to ℓ **do**
   1. Find the local pixel patch $\boldsymbol{P}_j^{(i)}$ for each sample $\boldsymbol{x}_j$
   2. Delete 10% outlier samples in $\boldsymbol{P}_j^{(i)}$
   3. Determine the intervals $[\boldsymbol{a}_j^{(i)}, \boldsymbol{b}_j^{(i)}]$ and box set $\boldsymbol{S}_j^{(i)}$
  end for
  4. Compute box kernel $K^{(i)}$ based on labeled pixels $\mathcal{L}$ and boxes $\boldsymbol{B}^{(i)}$ and train SVM model to obtain $\boldsymbol{\alpha}^{(i)}$
end for
Testing phase:
For any test sample $\hat{\boldsymbol{x}}$
for $i = 1$ to p
  1$'$. Perform the above Step **1**, **2**, **3** to obtain box $\hat{\boldsymbol{S}}^{(i)}$
  2$'$. Compute box kernel $\hat{K}^{(i)}$ based on $\hat{\boldsymbol{S}}^{(i)}$, $\mathcal{L}$ and $\boldsymbol{B}$
  3$'$. Compute multi-scale SVM outputs $\hat{f}^{(i)} = \hat{K}^{(i)}\boldsymbol{\alpha}^{(i)}$
end for
Output: The label as majority class in $\{\hat{f}^{(1)}, \dots, \hat{f}^{(p)}\}$

for the boundary pixels. Even if all points in the patch are homogeneous samples, deleting 10% outliers will not affect the intervals too much. Setting the lower and upper bound as 25-th and 75-th percentiles can be considered as another outlier elimination process. It eliminates the abrupt points in each band and helps to obtain a mild interval. It is no need to worry about these two outlier elimination processes will reduce the intervals because multi-scale patches are being used.

3) Effectiveness of classification to the case with limited training samples. In SVM-MPBK, all pixels in a box are intuitively assigned as the same class label. The labeled boxes can be considered as additional training samples. Therefore, compared with classic SVM methods, SVM-MPBK is more effective to the case with limited training samples.
4) Defining a local homogeneous region (or box) can be more flexible if we have more prior knowledge on the true spatial region distribution of each material. In this case, only a few boxes are needed. It is not necessary to build a box for each labeled pixel.
5) When fixing the patch scale, SVM-MPBK becomes SVM with a box kernel (SVM-BK) which can be considered as a generalization of SVM-CK. Different from SVM-CK that uses the mean (or standard deviation) of the pixels in a patch, SVM-BK uses the lower and upper bounds of the pixels in the patch to form a box. That is, SVM-CK operates on pixel points while SVM-BK operates on pixel sets. In the extreme case, the lower and upper bounds are chosen as the mean. Then the box set is reduced to the mean sample, and SVM-BK can revert SVM-CK.

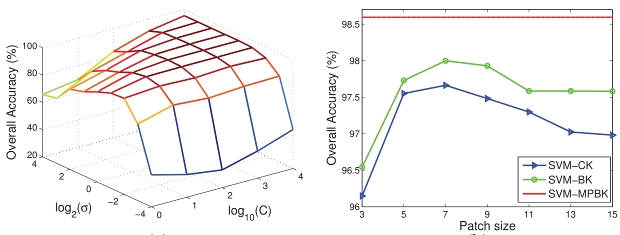
IV. Experimental Results

We demonstrate the effectiveness of the proposed SVM-MPBK on three benchmark hyperspectral data sets. SVM-MPBK is compared with SVM and SVM-CK. SVM uses only spectral information. SVM-CK uses both spectral and spatial features. The spatial feature is represented as the mean of pixels in a local pixel patch. The weighted summation kernel is used in SVM-CK. Gaussian kernel is used in all algorithms, and LIBSVM software is used to implement SVM [9].

A. Experiments on the Botswana Data Set

The first hyperspectral data is the Botswana data. This data was acquired by NASA EO-1 satellite over the Okavango Delta, Botswana in May 31, 2001 [10]. The image scene has the size of 1476×256 pixels. By discarding water absorption and noisy bands, 145 bands are retained. The data contain 3428 samples from 14 identified classes [10].[1]

The sensitivity of SVM-MPBK to the parameters is first discussed. Similar to classic SVM methods, SVM-MPBK has two tuning parameters, regularization parameter C and Gaussian kernel parameter σ. The regularization parameter C is chosen from the set $\{1, 10, 100, 1000, 10000\}$, and Gaussian kernel width σ varies in the range $\{2^{-4}, 2^{-3}, \dots, 2^{4}\}$. Fifteen samples in each class are selected as the training set for the parameter selection. Five-fold cross validation is used. We present the classification overall accuracies (OAs) with different parameter values and choose the optimal parameter values such that the OA is maximized. The overall accuracy is computed by the ratio of the number of correctly classified test samples to the number of all test samples. Fig. 1 (a) shows the overall accuracy of SVM-MPBK as a function of C and σ. We can see that the OA is stable when C varies from 100 to 10000 and σ changes from 2^{-1} to 2^{2}. We set $C = 1000$ and $\sigma = 1$ in the experiments.



(a) (b)

Fig. 1. Parameter setting for Botswana data set. (a) Overall accuracy versus C and σ; (b) Overall accuracy versus patch size.

Then, we investigate how the neighborhood patch size affects the box-kernel-based SVM methods. The patch size affects the computation of the intervals and box. We consider seven different patch windows, i.e., 3×3, 5×5, $\dots$, 15×15. Fig. 1(b) shows the changes of overall accuracies as a function of the patch sizes. Because the SVM-CK result is also affected

[1] http://www.csr.utexas.edu/hyperspectral/data/Botswana/.

by the neighborhood size, we report the corresponding results as well. As can be seen, the 7×7 patch provides the best OA for SVM-CK and SVM-BK classifiers consistently. In the following experiments for Botswana, the 7×7 patch is used for SVM-CK and SVM-BK. In SVM-MPBK, the ensemble of the multi-scale classifications outperforms the optimal results in the single patch scale. Different patches result in different intervals and box sets. These multi-scale box sets can well describe various HSI homogeneous regions. The ensemble learning makes full use of the complementary information in different patches, which achieves higher OAs.

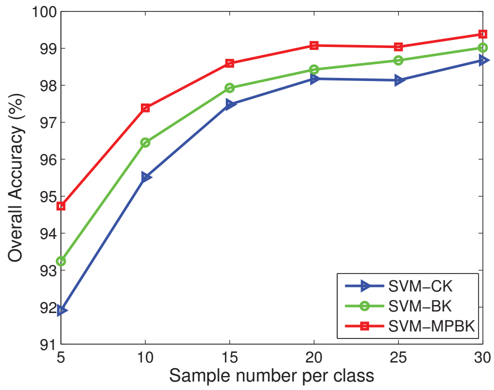


Fig. 2. OA versus sample number per class for Botswana data set.

TABLE I
CLASSIFICATION ACCURACIES (%) FOR BOTSWANA DATA SET.

Class	# Samples		Classification algorithms			
	Train	Test	SVM	SVM -CK	SVM -BK	SVM -MPBK
1	15	255	**99.84**	99.33	99.29	96.96
2	15	86	97.79	98.95	99.42	**100.0**
3	15	236	96.61	99.74	99.96	**100.0**
4	15	200	93.55	99.50	99.95	**100.0**
5	15	254	80.51	86.97	88.50	**92.95**
6	15	254	71.57	95.83	95.94	**97.44**
7	15	244	94.88	**99.71**	99.59	99.59
8	15	188	97.39	99.89	99.89	**100.0**
9	15	299	79.76	97.12	97.32	**97.93**
10	15	233	86.99	99.66	99.70	**99.74**
11	15	290	89.66	97.72	97.97	**98.07**
12	15	166	94.40	96.39	**99.82**	**99.82**
13	15	253	86.76	99.41	**100.0**	**100.0**
14	15	80	**97.25**	94.75	94.00	94.38
Overall accuracy			89.20	97.48	97.93	**98.59**
Average accuracy			90.50	97.50	97.95	**98.56**
Coefficient κ			0.883	0.973	0.978	**0.985**

Next, we show how the number of training samples affects the classification performance of SVM-CK, SVM-BK and SVM-MPBK. We randomly choose 5, 10, 15, 20, 25, 30 samples from each class to form the training set, respectively. The remaining samples are set as testing set. In each case, the experiment is repeated ten times with randomly chosen training samples. The results of ten runs are averaged. The classification overall accuracies under different number of training samples are shown in Fig. 2. With the increase of training samples, the overall accuracies for the three algorithms increase. SVM-BK improves SVM-CK because the box can capture local pixels variations more effectively than the mean vector. SVM-MPBK outperforms SVM-BK and SVM-CK consistently on different number of training samples.

In the following, we show the detailed classification performance of the algorithms in the case of 15 labeled samples per class. The classification overall accuracy, average accuracy, and coefficient κ on the test set are recorded. The average accuracy is the average of all class accuracies. Coefficient κ measures the degree of classification agreement. The number of training and testing samples for each class and the classification results are shown in Table I. It can be clearly seen that using spatial contextual information significantly improves classification performance (SVM versus SVM-CK), and modeling the local pixel correlation by means of box also benefits the classification (SVM-CK versus SVM-BK). By incorporating the complementary information in multi-scale patches, SVM-MPBK shows superior performance.

B. Experiments on the KSC Data Set

The Kennedy Space Center (KSC) data was acquired by NASA AVIRIS instrument over the Kennedy Space Center (KSC), Florida, on March 23, 1996 [10].[2] The image scene has 512×614 pixels and 224 spectral channels. By discarding water absorption and noisy bands, 176 bands are retained. It contains thirteen ground-truth classes. The total number of samples is 5211 ranging from 105 to 927 in each class.

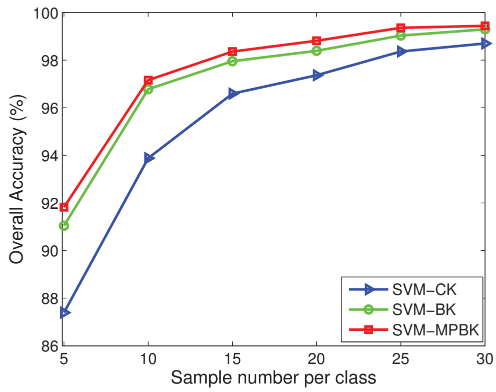


Fig. 3. OA versus sample number per class for KSC data set.

The classification overall accuracy of SVM-CK, SVM-BK and SVM-MPBK on different numbers of training samples are shown in Fig. 3. In SVM-CK, the neighborhood patch size is 11×11. In SVM-BK, $\sigma = 2$ and $C = 1000$ are used, and the neighborhood patch size is 13×13.

As can be seen from Fig. 3, when the number of training samples per class increases, the overall accuracies are greatly improved. SVM-BK shows a significant improvement over SVM-CK, especially in the cases of limited training samples.

[2] http://www.csr.utexas.edu/hyperspectral/data/KSC/.

The labeled boxes used in SVM-BK can be considered as additional training samples, thus SVM-BK can overcome the defect of small sample sized to some extent. SVM-MPBK slightly improves SVM-BK. When the number of training samples per class is 15, the classification overall accuracy, average accuracy, and coefficient κ on the test set are shown in Table II. Although the class accuracies of SVM-MPBK are not optimal in most of the classes, the overall accuracy, average accuracy, and coefficient κ of SVM-MPBK are consistently better than that of other methods.

TABLE II
CLASSIFICATION ACCURACIES (%) FOR KSC DATA SET.

Class	# Samples		Classification algorithms			
	Train	Test	SVM	SVM -CK	SVM -BK	SVM -MPBK
1	15	746	87.51	98.78	99.41	**99.50**
2	15	228	82.32	90.66	**95.26**	94.87
3	15	241	89.59	**99.21**	97.51	98.13
4	15	237	67.64	73.59	85.99	**92.83**
5	15	146	67.81	80.21	92.26	**93.77**
6	15	214	62.52	97.15	**99.44**	98.88
7	15	90	93.00	99.11	98.33	**99.33**
8	15	416	89.35	98.25	98.13	**98.80**
9	15	505	92.26	97.62	**98.71**	98.65
10	15	389	88.35	99.15	**99.36**	99.33
11	15	404	93.96	**98.49**	97.87	98.32
12	15	488	86.19	97.60	**98.20**	97.50
13	15	912	**100.0**	**100.0**	**100.0**	**100.0**
Overall accuracy			88.25	96.59	97.95	**98.36**
Average accuracy			84.65	94.60	96.96	**97.69**
Coefficient κ			0.869	0.962	0.977	**0.982**

C. Experiments on the Indian Pines Data Set

The Indian Pines data was acquired by the AVIRIS sensor in 1992.[3] The image scene contains with 145×145 pixels and 220 spectral bands, where 20 channels were discarded because of atmospheric affection. There are 16 classes in the data. The total number of samples is 10249 ranging from 20 to 2455 in each class.

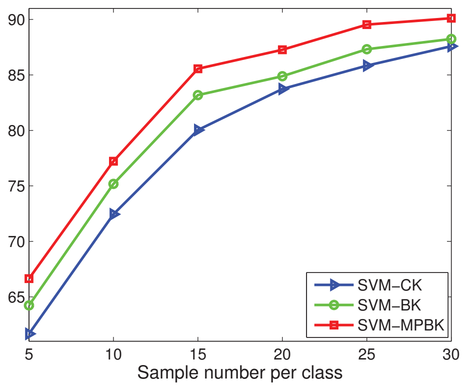


Fig. 4. OA versus sample number per class for Indian Pines data set.

[3] http://www.ehu.es/ccwintco/index.php/Hyperspectral_Remote_Sensing_Scenes.

TABLE III
CLASSIFICATION ACCURACIES (%) FOR INDIAN PINES DATA SET.

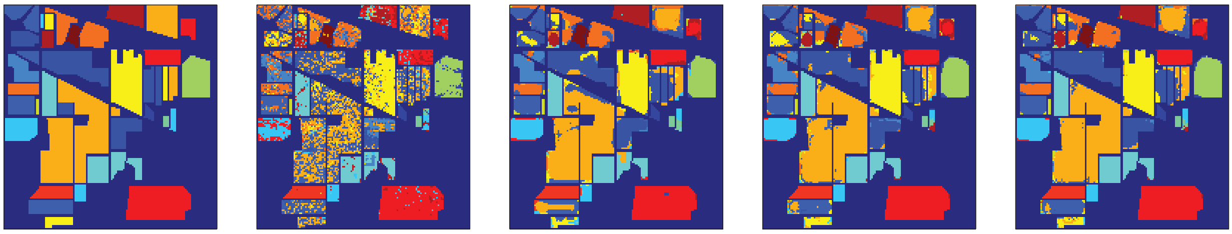
Class	# Samples		Classification algorithms			
	Train	Test	SVM	SVM -CK	SVM -BK	SVM -MPBK
1	15	31	88.71	**99.68**	98.06	**99.68**
2	15	1413	50.56	73.50	77.95	**81.90**
3	15	815	46.18	78.52	79.15	**84.56**
4	15	222	70.27	91.26	88.69	**92.79**
5	15	468	76.28	81.52	86.75	**88.31**
6	15	715	85.62	93.34	94.24	**96.06**
7	14	14	95.71	**100.0**	**100.0**	**100.0**
8	15	463	83.97	96.35	95.85	**97.34**
9	10	10	89.00	**100.0**	**100.0**	**100.0**
10	15	957	61.01	77.52	82.37	**85.29**
11	15	2440	46.61	69.86	76.30	**77.32**
12	15	578	45.24	75.24	76.76	**80.35**
13	15	190	96.00	**98.32**	95.00	96.63
14	15	1250	76.90	88.62	89.18	**90.67**
15	15	371	47.98	89.60	92.83	**94.23**
16	15	78	93.08	**98.21**	96.28	97.18
Overall accuracy			60.22	80.04	83.17	**85.55**
Average accuracy			72.07	88.22	89.34	**91.40**
Coefficient κ			0.553	0.775	0.810	**0.837**

Fig. 4 shows the classification overall accuracy on different numbers of training samples for the Indian Pines data set. For Grass-pasture-mowed and Oats classes with only few samples (28, 20 samples), at most half of total samples in the class are chosen for training. The neighborhood patch size is set as 15×15 for SVM-CK and 7×7 for SVM-BK. One can clearly see that SVM-BK provides higher overall accuracies than SVM-CK in all cases, and the superiority of SVM-BK is more evident when the number of training sample is very small. In the case of small sample size, the labeled boxes play a significant role via enlarging the training set. The additional information in the box helps to improve the generalization ability of classifier and provide a consistent classification results. When enough training samples are available, the information contained in the labeled pixels and labeled boxes tends to be consistent, and SVM-BK and SVM-CK will provide similar results. In all cases, SVM-MPBK outperforms SVM-BK and SVM-CK. This demonstrates that the combination of multi-scale boxes is more efficient in describing HSI homogeneous regions and the ensemble of complementary information in multi-scale outputs can significantly improve the classification performance.

The overall accuracy, average accuracy, and coefficient κ are shown in Table III. Fig. 5 shows the corresponding classification maps. The spectral-based SVM method produces classification results with more noises. Compared with other three methods, SVM-MPBK shows relatively better results.

V. CONCLUSION

In this paper, we have proposed a multi-scale patch based box kernel SVM method for classifying the local homogeneous regions of HSI. The proposed SVM-MPBK models the local homogeneous region of each pixel as a box and uses a box kernel to achieve simultaneous classification of labeled pixels and labeled boxes. Experimental results have shown



(a) (b) (c) (d) (e)

Fig. 5. Classification maps for Indian Pines data set. (a) Ground-truth. (b) SVM (OA=60.22%). (c) SVM-CK (OA=80.04%). (d) SVM-BK (OA=83.17%). (e) SVM-MPBK (OA=85.55%).

that, by integrating the labeled pixels and the prior knowledge on spatial homogeneous regions, the proposed SVM-MPBK outperforms spectral-based SVM and spatial-based SVM-CK.

ACKNOWLEDGMENT

This work was supported in part by the Macau Science and Technology Development Fund under Grant FDCT/106/2013/A3 and by the Research Committee at University of Macau under Grants MYRG2014-00003-FST, MRG017/ZYC/2014/FST, MYRG113(Y1-L3)-FST12-ZYC and MRG001/ZYC/2013/FST.

REFERENCES

[1] F. Melgani and L. Bruzzone, "Classification of hyperspectral remote sensing images with support vector machines," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 8, pp. 1778–1790, Aug. 2004.

[2] G. Camps-Valls and L. Bruzzone, "Kernel-based methods for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 6, pp. 1351–1362, Jun. 2005.

[3] G. Mountrakis, J. Im, and G. Ogole, "Support vector machines in remote sensing: A review," *ISPRS J. Photogramm. and Remote Sens.*, vol. 66, pp. 247–259, 2011.

[4] G. Camps-Valls, L. Gomez-Chova, J. Munoz-Maré, and et al., "Composite kernels for hyperspectral image classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 3, no. 1, pp. 93–97, 2006.

[5] A. Plaza, J. A. Benediktsson, J. W. Boardman, and et al., "Recent advances in techniques for hyperspectral image processing," *Remote Sens. Environ.*, vol. 113, pp. S110–S122, 2009.

[6] G. Camps-Valls, N. Shervashidze, and K. M. Borgwardt, "Spatio-spectral remote sensing image classification with graph kernels," *IEEE Geosci. Remote Sens. Lett.*, vol. 7, no. 4, pp. 741–745, Oct. 2010.

[7] G. Moser and S. B. Serpico, "Combining support vector machines and markov random fields in an integrated framework for contextual image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 5, pp. 2734–2752, May 2013.

[8] S. Melacci and M. Gori, "Learning with box kernels," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 11, pp. 2680–2692, Nov. 2013.

[9] C. Chang and C. Lin, "LIBSVM : A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, pp. 27:1–27:27, 2011. Software available at: http://www.csie.ntu.edu.tw/~cjlin/libsvm.

[10] J. Ham, Y. Chen, M. M. Crawford, and J. Ghosh, "Investigation of the random forest framework for classification of hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 3, pp. 492–501, Mar. 2005.