Spectral Shift Mitigation for Cross-Scene Hyperspectral Imagery Classification

Huan Liu, Wei Li, *Senior Member, IEEE*, Xiang-Gen Xia, *Fellow, IEEE*, Mengmeng Zhang, Chen-Zhong Gao, and Ran Tao, *Senior Member, IEEE*

***Abstract*—In cross-scene hyperspectral imagery (HSI) classification, labeled samples are only available in source scene, and how to properly reduce the spectral shift between source and target scenes is a matter of concern. In this article, we investigate this issue by considering the causes of the spectral shift and propose spectral shift mitigation (SSM) that includes amplitude shift mitigation (ASM) and adjacency effect mitigation (AEM). First, in ASM, the amplitude shift between source and target scenes is reduced by employing amplitude normalization on pixels of both source and target scenes. Then, in AEM, the spectral variation of target scene caused by adjacency effect is reduced by taking the weighted average spectral vector of surrounding pixels of a query pixel as the new spectral vector of the query pixel. Finally, a classifier trained by labeled samples from source scene is used for target scene. Superior classification performance on several cross-scene HSI data pairs demonstrates the effectiveness of the proposed SSM.**

***Index Terms*—Adjacency effect, amplitude normalization (SN), amplitude shift, cross-scene hyperspectral imagery classification, hyperspectral imagery (HSI).**

I. Introduction

HYPERSPECTRAL imagery (HSI) is an important source for land use and land cover classification in pixel level due to its abundant reflectance spectral information [1]–[13]. In HSI classification, to classify a scene correctly, sufficient labeled samples in the same scene are generally required. However, it is impractical and uneconomic to label new samples whenever classifying a new scene since labeling process is labor-consuming and time-costing. Meanwhile, different scenes may share same classes, and these labeled classes in source scene may share similar properties with the same classes in target scene. It is a nature to hope that the information from labeled samples in source scene can be exploited in target scene. This task is called cross-scene HSI classification, which aims to utilize the annotation information in source scene to help target scene classification.

One may consider directly using the labeled spectral signatures in source scene to train a classifier to classify target scene. This is a straightforward choice as spectral features in source and target scenes are in the same bands. However, spectral signatures belonging to the same land cover class may vary in different scenes, i.e., spectral shift. It is mainly caused by atmospheric and illumination condition, sensor nonlinearity, adjacency effect, seasonal variations, and variation in the material with time [14]–[20].

The problem of cross-scene HSI classification can be considered to be in the field of domain adaptation. For domain adaptation, many methods have been proposed in recent years and can be divided into three categories: Statistical alignment, subspace learning, and active learning. From the statistical perspective, the spectral shift is alleviated by reducing distribution difference between source and target scenes. To alleviate the distribution bias between source and target scenes, the usual way is to minimize the discrepancy of the statistics between source and target scenes. For the first-order statistic alignment, maximum mean discrepancy (MMD) criterion is a widely used technique [21]–[27]. Sun *et al.* [25] introduced multiple-kernel learning to simultaneously minimize MMD and the structural risk functional of support vector machines. Matasci *et al.* [26] introduced transfer component analysis (TCA) and semisupervised TCA [21] to cross-scene HSI classification and obtained acceptable cross-scene classification performance. By minimizing MMDs of marginal and conditional distributions simultaneously, Sun *et al.* proposed joint distribution adaptation (JDA) [24]. Peng *et al.* [28] proposed discriminative transfer joint matching (TJM) by embedding label information in TJM [29]. Xia *et al.* [27] proposed ensemble of TCA by incorporating TCA with a rotation-based ensemble strategy. For higher order statistics, Sun *et al.* [30] proposed CORAL by aligning correlation of source and target scenes. Ma *et al.* [31] proposed a method called class centroid and covariance alignment, where the first- and second-order statistics are aligned on a per-class basis.

From the subspace learning perspective, source and target scenes have a common subspace where labeled samples of source scene are sufficiently informative to discriminate concerned classes in target scene. Most subspace learning methods learn the common subspace by aligning the geometrical structures of source and target scenes. Ye *et al.* [32] proposed

Manuscript received May 8, 2021; revised May 28, 2021; accepted June 20, 2021. Date of publication June 22, 2021; date of current version July 14, 2021. This work was supported in part by the Bejing Natural Science Foundation under Grant JQ20021, and in part by the National Natural Science Foundation of China under Grant 61922013 and Grant U1833203. *(Corresponding author: Wei Li.)*

Xiang-Gen Xia is with the School of Information, and Electronics, Beijing Institute of Technology, Beijing 100081, China, and also with the Department of Electrical, and Computer Engineering, University of Delaware, Newark, DE 19716 USA (e-mail: xianggen@udel.edu).

Huan Liu, Wei Li, Mengmeng Zhang, Chen-Zhong Gao, and Ran Tao are with the School of Information, and Electronics, Beijing Institute of Technology, Beijing 100081, China (e-mail: huanliu233@gmail.com; leewei36@gmail.com; 7520200002@bit.edu.cn; 3120205425@bit.edu.cn; rantao@bit.edu.cn).

Digital Object Identifier 10.1109/JSTARS.2021.3091591

multitask joint dictionary learning scheme to learn a common dictionary for source and target scenes in feature extraction level, and proposed a multitask sparse logistic regression (SRL) in classifier level to further improve cross-scene classification performance. Samat *et al.* [33] introduced geodesic flow kernel (GFK) [34] to cross-scene HSI classification, where an infinite number of subspaces that characterize changes from source scene to target scene are integrated. By combining GFK with multiple kernel learning and spatial features–extended multi-attribute profiles (EMAPs), spectral-spatial multiple geodesic flow kernel learning (S2-MGFKL) was proposed in [35]. Some subspace learning methods are more complex. Stratified transfer learning (STL) [36] learns the common subspace on a per-class basis by exploiting the intraaffinity of classes basis, and the strategy of major voting of different classifiers (KNN, SVM, and random forest) is also utilized. By minimizing distribution shift and subspace shift simultaneously, joint geometrical and statistical alignment (JGSA) was proposed in [37]. Manifold embedded distribution alignment (MEDA) was proposed to learn the common subspace by incorporating manifold constraint into distribution alignment [38]. Ghifary *et al.* proposed scatter component analysis (SCA) to learn the common subspace by minimizing between-domain and within-class scatter, and maximizing total scatter and between-class scatter simultaneously [39]. From the active learning perspective, based on the classifier initially trained by labeled samples in source scene, samples in target scene are selected iteratively as new training samples to improve cross-scene classification performance. In [40], two strategies based on uncertainty and clustering of the data space were introduced to perform sample selection under their corresponding heuristics. In [41], two queries were created. One is to select samples in target scene as new training samples that minimize the difference between the largest and the second largest class-conditional densities. The other is to delete labeled samples in source scene that maximize the difference between the initial class-conditional densities trained only by labeled samples in source scene and the current class-conditional densities.

In recent years, deep learning-based domain adaptation methods have been attracting a lot attention, and they combine the techniques used in the traditional methods with deep artificial neural networks. Deep adaptation network (DAN) used multiple kernel MMD (MK-MMD) to align the deep features of source and target domains [42]. By combining CORAL with neural network, deep-CORAL was developed [30]. More recently, deep subdomain adaption network (DSAN) was proposed by aligning the newly defined subdomains with local MMD (LMMD), where pseudosoft label information was utilized [43].

However, domain adaptation methods of distribution alignment, subspace learning, and active learning mentioned above blindly align source and target domains without considering the causes of the discrepancy between source and target domains. This issue also arises in learning-based domain adaptation neural networks mentioned above although deep features can be learned from these networks. As a result, some problems may occur in cross-scene HSI classification. For example, if no spectral shift exists between source and target scenes, labeled samples in source scene can be well utilized to help target scene classification, and any domain adaptation technique seems unnecessary. However, if label distributions of source and target scenes are completely different, the marginal distributions of source and target scenes are naturally different. For widely MMD-based methods, this discrepancy should be narrowed. As a result, the marginal distributions of the transformed features in source and target scenes become similar, but to do so the transformed features of one class in source and target scenes become different inevitably. Let us take a simple example: Scene $\mathcal{A}$ has 100 red and 10 blue balls, while scene $\mathcal{B}$ has 10 red and 100 blue balls. To make the marginal distribution of scene $\mathcal{B}$ similar to that of scene $\mathcal{A}$, a straightforward transformation is to take red balls in scene $\mathcal{B}$ as "blue" while blue balls as "red." Consequently, the intrinsic colors of scene $\mathcal{B}$ have changed, i.e., the meanings of one color (red or blue) in scenes $\mathcal{A}$ and $\mathcal{B}$ after transformation become different. For cross-scene HSI classification, this transformation induces new "spectral shift," and thus, deteriorates cross-scene HSI classification performance.

In this article, by tackling spectral shifts directly, which are the causes of the discrepancy between source and target scenes, a method called spectral shift mitigation (SSM) is developed, including amplitude shift mitigation (ASM) and adjacency effect mitigation (AEM). By conducting some spectral imaging experiments on man-made materials, and studying some public HSI datasets, it is observed that the shape of spectral curve of the same class is stable but the amplitude varies considerably over scenes. In ASM, amplitude normalization (AN) with that the spectral vector is normalized to have unit l_1-norm, is proposed to mitigate the amplitude shift. Furthermore, since a hyperspectral sensor gathers not only the directly reflective power from a query pixel but also the indirectly diffuse powers from its adjacent pixels in the scene, the spectral curve of the query pixel is affected by these adjacent pixels, i.e., the spectral shift or variation caused by adjacency effect [18]. This considerable variation of spectral curves in target scene may degrade cross-scene classification performance. In AEM, the weighted average spectral vector of surrounding pixels to a query pixel with affinity coefficients as their weights, is taken as the new spectral vector of the query pixel.

The main contributions of this article are listed as follows: 1) By reducing spectral shifts directly, two strategies are proposed to improve the performance of cross-scene HSI classification. One strategy is to reduce the amplitude shift, and the other is to reduce the spectral shift on target scene caused by adjacency effect. 2) Based on the first proposed strategy, ASM is developed to reduce amplitude shift that may be a large spectral shift between source and target scenes. 3) Based on the second strategy, AEM is implemented on target scene to reduce the spectral shift caused by adjacency effect.

The remainder of this article is organized as follows. Section II shows the observation on some spectral imaging experiments and public HSI datasets. Section III explicitly explains the proposed SSM. Experimental results are shown in Section IV to demonstrate the superior performance of SSM. In Section V, we conclude this article.

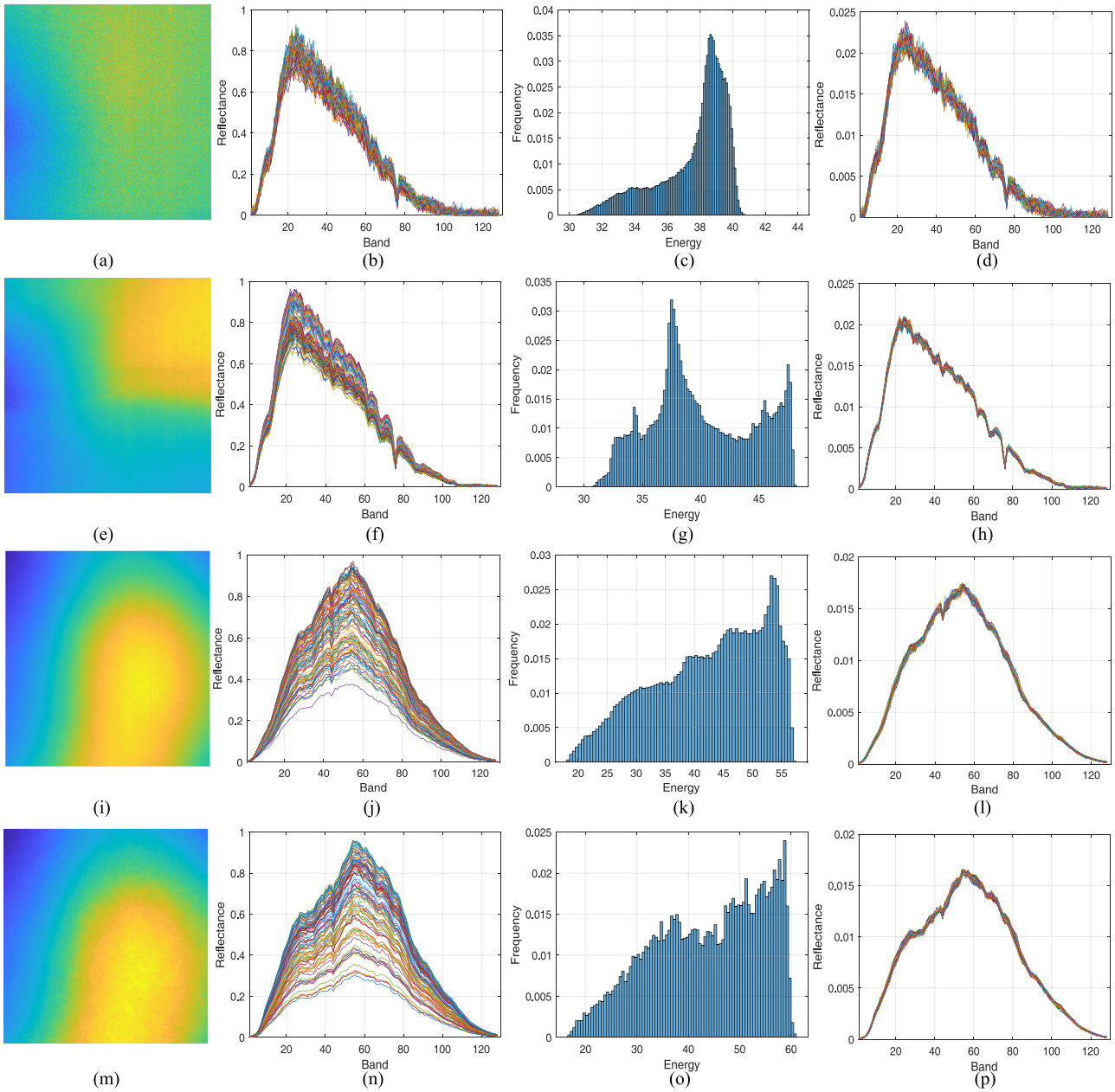

Fig. 1. Pseudocolor maps, spectral curves without AN, amplitude histograms and spectral curves with AN for different man-made materials in different illumination conditions: (a), (b), (c), (d) Reference panel illuminated by sunlight; (e), (f), (g), (h) Reference panel illuminated by halogen lamp; (i), (j), (k), (l) White paper illuminated by sunlight; (m), (n), (o), (p) White paper illuminated by halogen lamp.

II. Amplitude Shift in HSI

In this section, we conduct experimental observations on four HSI datasets acquired by our group and on one public cross-scene HSI data pair. From the experiments one can see that the shape of spectral curve of one material is stable while the amplitude varies dramatically over scenes. This phenomenon is defined as amplitude shift.

A. Amplitude Shift in Man-Made Materials

The four HSI datasets are reference panel by sunlight, reference panel by halogen lamp, white paper by sunlight, and white paper by halogen lamp. They are all acquired by SOC-170 Hyperspectral Imaging System indoor (by halogen lamp), or outdoor (by sunlight) by our group. Note that the incident flux of sunlight or halogen lamp is guaranteed to be almost invariant over all pixels. Reference panel is one attachment to SOC-170, and white paper is the usual paper for printing. They all consist of 656 $\times$ 520 pixels and 128 spectral bands covering 0.43–1.20 μm. The pseudocolor images of the four datasets are shown in Fig. 1(a), (e), (i), (m), respectively, and generated by the images of bands 13, 35, and 64. The spectral curves of the four datasets are shown in Fig. 1(b), (f), (j), (n), respectively, where spectral curves of 100 random samples instead of all samples are depicted to save storage.

Since every part of reference panel (or white paper) is made of the same elements, all pixels of reference panel (or white paper) are considered as one material. However, from Fig. 1(a), (e), (i), (m), the colors of pixels vary considerably, and pixels in different colors seem to represent different materials. Furthermore, as shown in Fig. 1(b), (f), (j), (n), the corresponding spectral curves of pixels of reference panel or white paper vary dramatically, and it is hard to recognize them as one material. Let us define amplitude of a pixel as the integral (continuous spectral curve) or sum (discrete spectral curve) of the spectral curve of a pixel. Divided by their amplitudes, i.e., under AN as shown in Fig. 1(d), (h), (l), (p), the spectral curves of all pixels in each dataset become almost consistent and indistinguishable. This means that the variation of the original spectral curve of one material can be absorbed by AN. From the histograms generated by using all pixels of the corresponding four datasets as shown in Fig. 1(c), (g), (k), (o), the amplitudes of pixels vary considerably. This shows that the shape of spectral curve of one material is nearly invariant but the amplitude varies considerably. Moreover, from Fig. 1(b), (f), (j), (n), not only spectral curves but also amplitude histograms of one material under different illumination conditions (sunlight and halogen lamp) are noticeably different, but with AN to reduce the amplitude shift, spectral curves of the same material become similar. This means that the distributions of amplitude of one material under different illumination conditions may be different. This kind of spectral shift is defined as amplitude shift. By reducing amplitude shift, we can not only reduce the difference of pixels of one material under the same illumination condition but also reduce the discrepancy of pixels of the same material under different illumination conditions. Therefore, reducing amplitude shift in source and target scenes seems to be important for cross-scene HSI classification. Any normalization that can reduce amplitude shift can help obtain more invariant features. For example, besides AN, length normalization (LN) that normalizes a spectral vector to have unit l_2-norm is a good normalization method for reducing amplitude shift as well.

B. Amplitude Shift in Public HSI Datasets

In last subsection, man-made materials are studied to show the amplitude shift of one material. Here, public HSI datasets are studied to further confirm the observation. Houston data pair is chosen and it consists of Houston2013 and Houston2018 datasets [44], [45]. Houston2013 was gathered by the CASI-1500 sensor over the University of Houston campus and neighboring areas in June 2012. It consists of 349×1905 pixels with spatial resolution of 2.5 m and 144 spectral bands in the wavelength range 0.38-1.05 μm. Houston2018 was captured by the ITRES CASI-1500 sensor over the University of Houston campus in February 2017. It consists of 601×2384 pixels with spatial resolution of 1 m and 48 spectral bands in the wavelength range $0.38 - 1.05$ μm. To make Houston2013 consistent with Houston2018, seven classes are chosen, 48 bands are extracted, and an overlapping area of 210×954 with Houston2018 is selected. The labeled samples of Houston2013 and Houston2018 are shown in Table I.

TABLE I
CLASS DISTRIBUTION OF SAMPLES FOR HOUSTON DATA PAIR

No.	Class Name	Houston2013	Houston2018
1	Grass healthy	345	1353
2	Grass stressed	365	4888
3	Trees	365	2766
4	Water	285	22
5	Residential buildings	319	5347
6	Non-Residential buildings	408	32459
7	Road	443	6365
Total		2530	53200

Considerable amplitude variation exists in one scene as depicted in Fig. 2(a). There are two kinds of spectral curves of water in Houston2018 dataset, and they are appreciably different. From Fig. 2(b), after performing AN, their difference is significantly narrowed. From Fig. 3(a), amplitude variation is more considerable between different scenes. Mean spectral curves of water in Houston2013 and Houston2018 datasets are significantly different due to amplitude shift. After performing AN, the difference is also narrowed visibly in Fig. 3(b). Comparing amplitude histograms of Houston2013 and Houston2018 datasets in Fig. 3(c) and (d), they are noticeably different. It can be expected that better cross-scene classification performance can be obtained when amplitude shift is alleviated.

C. Analysis of Amplitude Shift

Although many causes of spectral shift have been described in the Introduction, the only cause that may be related to amplitude shift is illumination condition, where amplitudes of sunlight in different parts of one material are different. As is known, the spectral curve of a pixel is reflectance curve. The reflectance R is defined as the ratio of reflected flux F_r to incident flux F_i to the pixel [46], i.e.,

$$R(\lambda) = \frac{F_r(\lambda)}{F_i(\lambda)} \tag{1}$$

where λ is wavelength and $R(\lambda)$ is wavelength-related. It may be impractical and uneconomic to measure incident flux F_i of every pixel. Generally, for convenience we may assume that incident flux F_i is invariant for every pixel. However, in fact, incident flux F_i may vary over pixels. Due to the neglect of variation of incident flux to different pixels, the reflection curves of one material may be different over pixels, resulting in amplitude shift.

However, the amplitude shift brought by the difference of incident flux can be actually ignored in some situations. For example, for remote sensing images, the incident flux to all pixels can be regarded as equal since the source of incident light is sunlight. Another example is the illumination condition of the four HSI datasets gathered by our group. For every pixel, the incident flux of sunlight or halogen lamp is guaranteed to be almost invariant over pixels, but the phenomenon of amplitude shift is still present in all four datasets. Therefore, a primary cause for amplitude shift should be explored. This primary cause is the geometrical structure of the surface of a material, and more precisely, the smoothness of the surface.

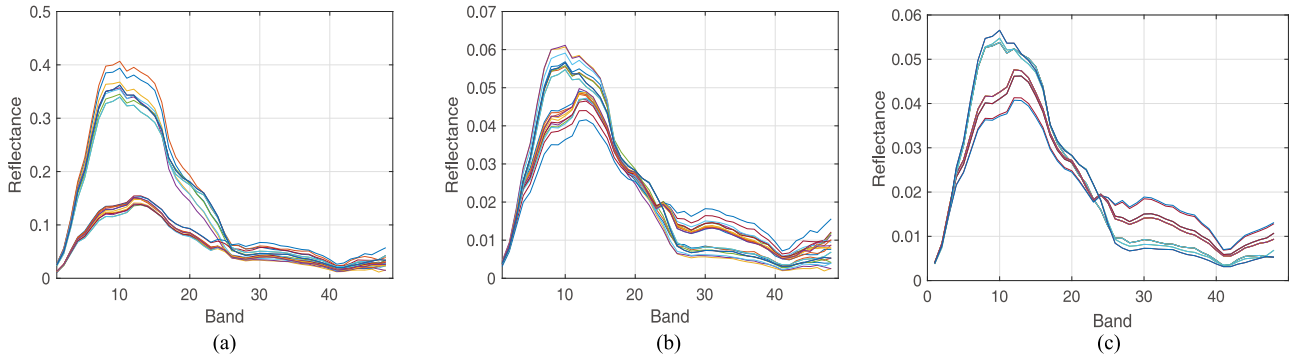

Fig. 2. Spectral curves of water in Houston2018 dataset. (a) Spectral curves of water in original Houston2018 dataset. (b) Spectral curves of water with AN in Houston2018 dataset. (c) Spectral curves of water with AN in Houston2018 dataset after performing AEM.

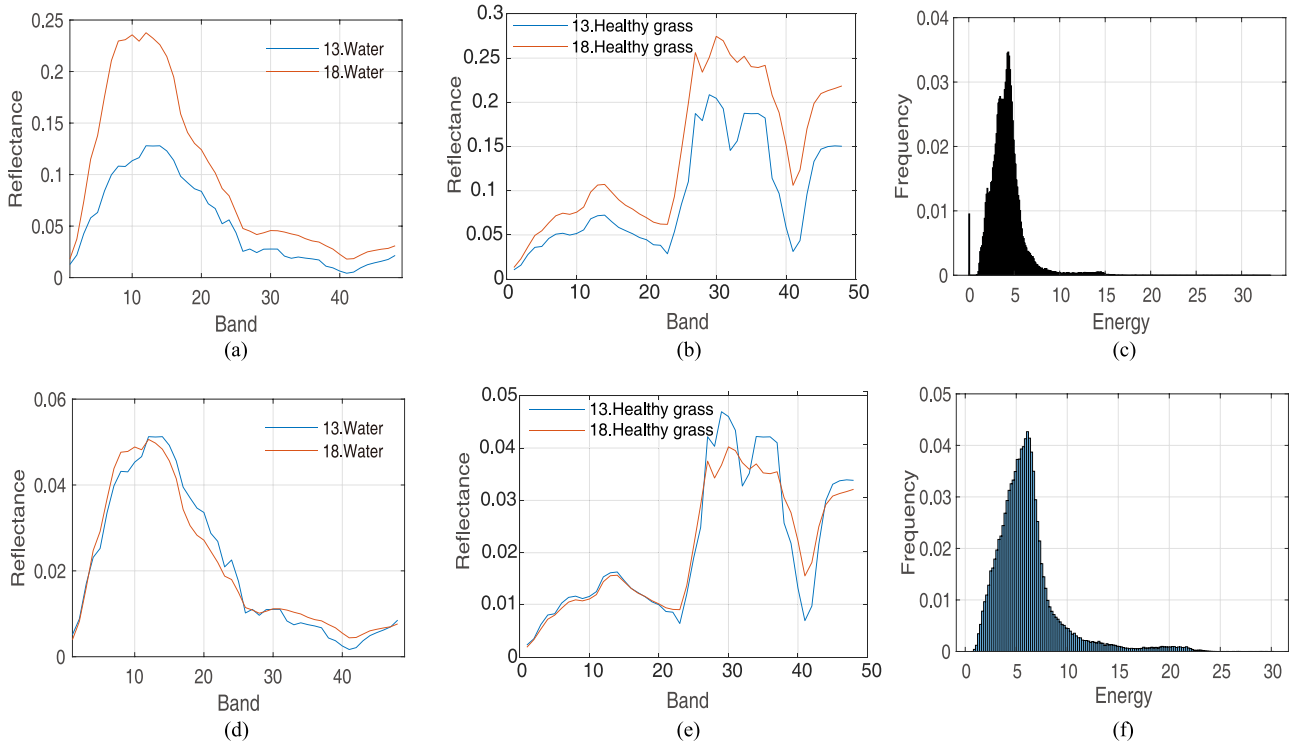

Fig. 3. Comparisons between Houston2013 and Houston2018 datasets. (a), (b) Mean spectral curves of water without and with AN in Houston2013 and Houston2018 datasets. (c), (d) Mean spectral curves of healthy grass without and with AN in Houston2013 and Houston2018 datasets. (e), (f) Amplitude histograms of Houston2013 and Houston2018 datasets, respectively.

Suppose that there are two panels with the same material, but the surface of one panel is smoother than that of the other. Since the reflection of the smoother panel is closer to mirror reflection while that of the rougher one is closer to diffuse reflection, the average of image intensity of the smoother panel is greater than that of the rougher one, though illuminated by the light of the same amplitude. By (1), the reflectance of the smoother panel is greater than that of the rougher one. Since they are made of the same material, the ratio of reflectance of the smoother panel to that of the rougher panel should be the relative smoothness of the smoother panel to the rougher panel, and it is independent of wavelength λ. As a result, amplitude shift happens between pixels of one class with different smoothness.

III. Proposed SSM Framework

In this section, the motivation for the proposed SSM is introduced. Then, the flowchart of the proposed SSM is described, including ASM and AEM processes as shown Fig. 4.

A. Motivation

As pointed out in the Introduction, the difficulty of cross-scene HSI classification lies in the spectral shift problem caused by many factors. However, the spectral shift caused by atmospheric and illumination condition, and sensor nonlinearity is commonly alleviated in the preprocessing stage. The first step is to perform radiometric correction for at-sensor reflectance.

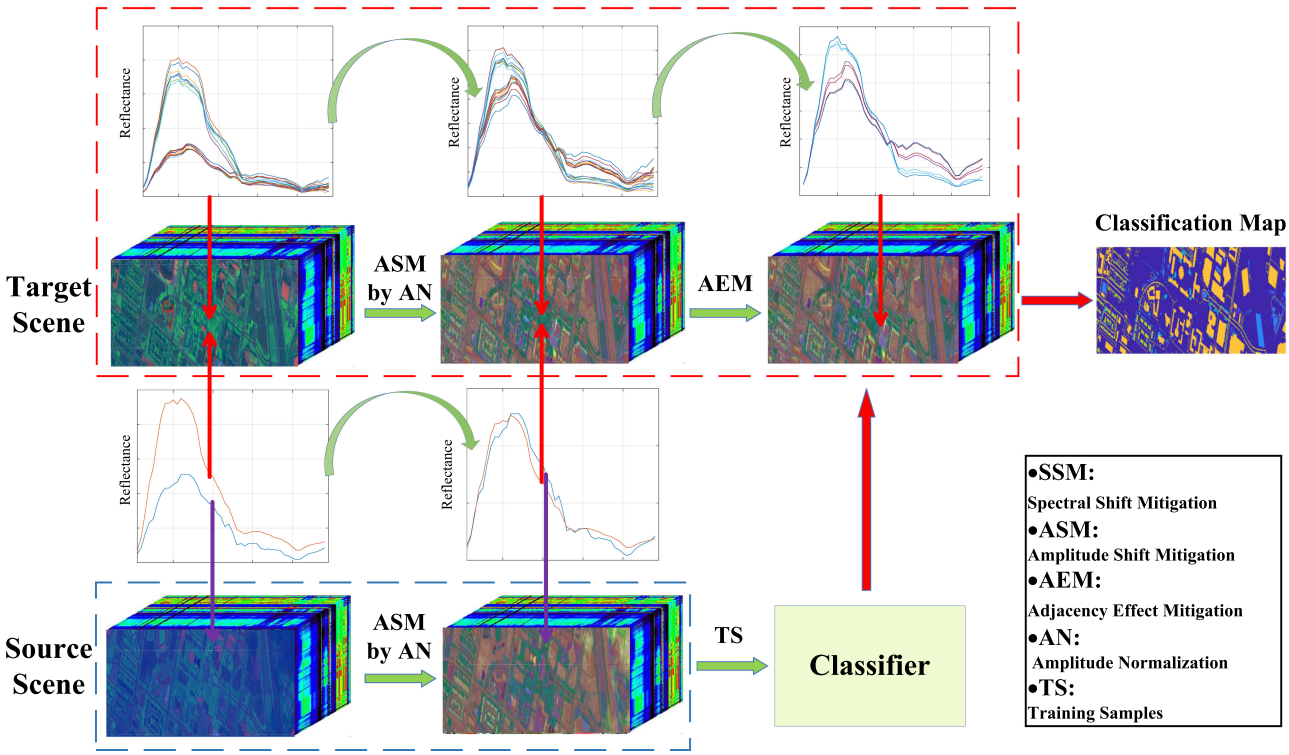

Fig. 4. Flowchart of the proposed SSM framework, including ASM and AEM processes.

To correct for atmospheric degradations, illumination effects, and sensor response differences, a scene-to-scene radiometric normalization technique was proposed in [14]. BRDF effects correction (BREFCOR) is another method for radiometric correction [47]. It improves the consistency of surface reflectance products and is applicable for various kinds of imaging spectrometer systems and surface characteristics. In the second step, at-sensor reflectance values should be converted to the bottom of atmosphere reflectance values. This is commonly done by some methods, such as ATCOR-4 [48] and MODTRAN [49]. In general, original HSI data gathered by sensors has been already processed by some of these methods before it is used in subsequent HSI-based tasks, including HSI classification. By these methods, the discrepancy between source and target scenes has been already reduced to some extent.

Although these methods have alleviated the spectral shift caused by atmospheric and illumination condition, and sensor nonlinearity in the preprocessing stage, one spectral shift, i.e., amplitude shift caused by the geometrical structure of surface, has not been alleviated. As a result, the amplitude distributions of source and target scenes are quite different. To improve cross-scene HSI classification performance, amplitude shift has to be alleviated. The proposed SSM is a method to mitigate the amplitude shift caused by geometrical structure of surface, and consequently the discrepancy between source and target scenes is further reduced.

Furthermore, the spectral shift caused by adjacency effect as explained in the Introduction should be mitigated as well since considerable spectral variation of target scene may degrade cross-scene classification performance. By reducing the spectral variation of one class over pixels in target scene, the proposed SSM is a method to alleviate the spectral shift caused by adjacency effect so that better cross-scene classification performance can be obtained.

B. *Proposed SSM*

HSI is a 3D tensor $\mathbf{H} \in \mathbb{R}_+^{m\times n\times d}$, where $\mathbb{R}_+$ is the set of real numbers that are larger than or equal to zero, m and n are the height and width, respectively, and d is the number of spectral bands. For pixel-wise hyperspectral classification, we convert the 3D image into pixel-wise 2D matrix $\mathbf{X} = [\mathbf{x}_i] \in \mathbb{R}_+^{d\times mn}$ with $\mathbf{x}_i \in \mathbb{R}_+^d$. In the last section, we described a spectral shift–amplitude shift that the shape of hyperspectral curve of one material is nearly invariant but the amplitude varies considerably. Another spectral shift is that the spectral curve of a pixel of one class is affected by nearby pixels of different classes in the scene, i.e., the spectral shift caused by adjacency effect. The proposed SSM that includes ASM and AEM processes is to mitigate the two kinds of spectral shifts.

For the amplitude shift, if $\mathbf{x}_i$ belongs to one class, $\hat{\mathbf{x}}_i = a\mathbf{x}_i$ belongs to the same class with a constant $a > 0$ as the amplitude shift coefficient. To reduce the amplitude shift for cross-scene HSI classification, LN and AN can be applied as we shall see below. Note that LN was used in some works of HSI classification [3], [5], where they obtained satisfying classification performance, but its effect to reduce amplitude shift was not directly mentioned. AN is proposed here for the first time to reduce amplitude shift for cross-scene HSI classification. Suppose the normalized data are $\mathbf{Y} = [\mathbf{y}_i] \in \mathbb{R}_+^{d\times mn}$ with $\mathbf{y}_i \in \mathbb{R}_+^d$. For LN,

Algorithm 1: Proposed SSM Algorithm.

Input: Source data $\mathbf{S} = \{\mathbf{s}_j\}_{j=1}^N \in \mathbb{R}_+^d$, target data $\mathbf{X} = \{\mathbf{x}_j\}_{j=1}^N \in \mathbb{R}_+^d$, neighborhood radius w, iterative times n_t of AEM on target scene.
Step 1: Performing AN according to (5) on pixels of source and target scenes, and the results of source and target scenes are $\mathbf{S}' = \{\mathbf{s}'_j\}_{j=1}^N \in \mathbb{R}_+^d$ and $\mathbf{Y} = \{\mathbf{y}_j\}_{j=1}^N \in \mathbb{R}_+^d$, respectively;
Step 2: Initialing $\mathbf{Z}^{(0)} = \mathbf{Y} = \{\mathbf{y}_j\}_{j=1}^N$;
For $i = 1$ to n_t
Step 4: $\mathbf{Y} = \mathbf{Z}^{(i-1)}$;
Step 5: $\mathbf{Z}^{(i)} = \{\mathbf{z}_j\}_{j=1}^N$is obtained according to (6)–(9);
End For;
Step 6: $\mathbf{Z} = \mathbf{Z}^{(n_t)}$;
Output: The processed source and target data $\mathbf{S}'$ and $\mathbf{Z}$, respectively.

a spectral vector is normalized to have unit l_2-norm, i.e.,

$$\mathbf{y}_i = \frac{\mathbf{x}_i}{||\mathbf{x}_i||_2} = \frac{\mathbf{x}_i}{\sqrt{\sum_{b=1}^d |x_{ib}|^2}} \tag{2}$$

where x_{ib} is the reflectance of band b. Clearly, after performing LN, $\hat{\mathbf{x}}_i$ as amplitude shift version of $\mathbf{x}_i$ becomes

$$\begin{aligned}\mathbf{y}'_i &= \frac{\hat{\mathbf{x}}_i}{||\hat{\mathbf{x}}_i||_2} = \frac{\hat{\mathbf{x}}_i}{\sqrt{\sum_{b=1}^d |\hat{x}_{ib}|^2}} \\ &= \frac{a\mathbf{x}_i}{\sqrt{\sum_{b=1}^d |ax_{ib}|^2}} \\ &= \frac{\mathbf{x}_i}{\sqrt{\sum_{b=1}^d |x_{ib}|^2}} = \mathbf{y}_i\end{aligned} \tag{3}$$

where amplitude shift a is eliminated by LN. For AN, a spectral vector is normalized to have unit l_1-norm, i.e.,

$$\mathbf{y}_i = \frac{\mathbf{x}_i}{||\mathbf{x}_i||_1} = \frac{\mathbf{x}_i}{\sum_{b=1}^d |x_{ib}|}. \tag{4}$$

Similar to LN, after performing AN, $\hat{\mathbf{x}}_i$ as amplitude shift version of $\mathbf{x}_i$ becomes

$$\begin{aligned}\mathbf{y}'_i &= \frac{\hat{\mathbf{x}}_i}{||\hat{\mathbf{x}}_i||_1} = \frac{\hat{\mathbf{x}}_i}{\sum_{b=1}^d |\hat{x}_{ib}|} \\ &= \frac{a\mathbf{x}_i}{\sum_{b=1}^d |ax_{ib}|} \\ &= \frac{\mathbf{x}_i}{\sum_{b=1}^d |x_{ib}|} = \mathbf{y}_i\end{aligned} \tag{5}$$

where amplitude shift a is eliminated by AN as well. For the amplitude shift mitigation, instead of the above LN and AN, using other norms is fine too. However, from the following, we will see that AN is unique for our proposed SSM algorithm. Also note that the above LN and AN may apply not only to nonnegative valued input $\mathbf{x}_i$ but also to general complex-valued input $\mathbf{x}_i$.

For the adjacency effect, the spectral vector of a query pixel of one class is affected by its surrounding pixels. These surrounding pixels include pixels of the same class and pixels of different classes. The pixels from different classes generate spectral shift for the query pixel while the pixels from the same class help reduce spectral shift. Obviously, to reduce the spectral shift caused adjacency effect, the effect of pixels of the same class should be augmented while that of different classes should be reduced. One natural way to realize this is to take the average spectral curve of the surrounding pixels of the same class as the new spectral vector of the query pixel. However, this method is unrealizable since the classes of the surrounding pixels are not provided. One way to deal with this problem is to take the weighted average spectral vector of all the surrounding pixels as the new spectral vector of the query pixel, and the corresponding weights are the affinity coefficients of the surrounding pixels with the query pixel. This is based on the fact that adjacent pixels with large affinity coefficients belong to the same class with high probability. As a consequence, the effect of surrounding pixels that are more likely to be in the same class with the query pixel is augmented. In AEM, for a query pixel $\mathbf{y}_i$, the new spectral vector is

$$\mathbf{z}_i = \sum_{\mathbf{y}_j \in N_w(\mathbf{y}_i)} w_j \mathbf{y}_j \tag{6}$$

where $N_w(\mathbf{y}_i)$ is a set of spatially adjacent pixels of $\mathbf{y}_i$ with neighborhood radius of w, and the weight or affinity coefficient w_j of $\mathbf{y}_j$ belonging to $N_w(\mathbf{y}_i)$ satisfies

$$\sum_j w_j = 1 \tag{7}$$

and

$$w_j = \frac{1/d_j}{\sum_{\mathbf{y}_k \in N_w(\mathbf{y}_i)} 1/d_k} \tag{8}$$

with

$$d_j = \sqrt{\sum_{b=1}^d |y_{ib} - y_{jb}|^2}. \tag{9}$$

From (6), $\mathbf{z}_i$ is closer to the surrounding pixels with higher affinity coefficients since they have more weights. Therefore, AEM makes the query pixel closer to the surrounding pixels that are more likely to be in the same class with query pixel. However, a satisfying result may not be obtained by performing AEM for one time. To reduce the spectral shift caused by adjacency effect more effectively, we can perform AEM iteratively for n_t times, i.e., performing AEM again on the last result of AEM repeatedly for n_t times. Consequently, the query pixel is closer and closer to the surrounding pixels that are more likely to be in the same class during the iteration, and the effect of these surrounding pixels is further augmented. To show the effect of AEM to reduce the variation of target, spectral curves of water in Houston2018 dataset before and after performing AEM are depicted in Fig. 3 (b) and (c), respectively, where several curves cluster into one curve, significantly reducing the variation of spectral curves of water in Houston2018 dataset.

From (8) and (9), if ASM is not implemented first, affinity coefficients calculated by the original data cannot truly reflect the affinities between the query pixel and its surrounding pixels due to amplitude shift. Therefore, ASM has to be implemented first to ensure a satisfying calculation in affinity coefficients. In ASM, a suitable normalization method that incorporates AEM needs to be considered. It is natural to require that the normalized data after implementing AEM should still remain normalized. For example, according to this requirement, length-normalized data after implementing AEM should still have unit l_2-norm. However, according to the triangular inequality

$$\begin{aligned} ||\mathbf{z}_i||_2 &= || \sum_{\mathbf{y}_j \in N_w(\mathbf{y}_i)} w_j \mathbf{y}_j ||_2 \\ &\leq \sum_{\mathbf{y}_j \in N_w(\mathbf{y}_i)} w_j ||\mathbf{y}_j||_2 \\ &= \sum_j w_j = 1 \end{aligned} \tag{10}$$

where the equality can be obtained only when for all vectors $\mathbf{y}_j$ in the neighborhood $N_w(\mathbf{y}_i)$, all vectors $w_j \mathbf{y}_j$ have the same direction. This clearly may not hold in general. Thus, LN cannot meet the requirement. On the contrary, AN satisfies this requirement naturally as the l_1-norm of the normalized data with AN after implementing AEM is still unit, i.e.,

$$\begin{aligned} ||\mathbf{z}_i||_1 &= || \sum_{\mathbf{y}_j \in N_w(\mathbf{y}_i)} w_j \mathbf{y}_j ||_1 \\ &= \sum_{b=1}^{d} \left| \sum_{\mathbf{y}_j \in N_w(\mathbf{y}_i)} w_j y_{jb} \right| \\ &= \sum_{\mathbf{y}_j \in N_w(\mathbf{y}_i)} w_j ||\mathbf{y}_j||_1 \\ &= \sum_j w_j = 1 \end{aligned} \tag{11}$$

where the second equality is based on the fact $w_j,\ y_{jb} > 0$. Therefore, AN is a natural choice for ASM. In fact, from the above analysis it is not hard to see that our newly proposed AN is the only choice to keep the the unit norm after AEM among other commonly used norms.

The overall description of the proposed SSM is given as Algorithm 1. First, ASM by AN is exerted on both source and target scenes to reduce the spectral shift caused by amplitude shift between source and target scenes. Then, AEM is performed on target scene to further reduce the spectral shift caused by adjacency effect. Finally, a classifier trained by labeled samples in source scene is used for target scene classification.

IV. Experimental Results and Analysis

In this section, the proposed SSM is utilized on two cross-scene HSI data pairs to validate its effectiveness in cross-scene HSI classification. First, these two HSI data pairs are introduced, including Houston data pair, and Shanghai–Hangzhou data pair. Second, the impact of neighborhood radius w and iterative times n_t of AEM on the cross-scene classification performance is quantified by the classical SVM classifier [1]. For each data pair, one dataset can be used as source, and the other can be used as target. Then, we have four cases of cross-scene classification, i.e., Houston2013→Houston2018, Houston2018→Houston2018, Shanghai → Hangzhou, and Hangzhou → Shanghai. Finally, based on SVM classifier, experiments of SSM on the four cases of cross-scene classification are carried out in comparison with several traditional and state-of-the-art algorithms for cross-scene classification, including traditional SVM, CORAL [30], TCA [21], JDA [24], JGSA [37], MEDA [38], SCA [39], STL [36], and DSAN [43]. The class-specific accuracy (CA), overall accuracy (OA), average accuracy, and Kappa coefficient are employed to evaluate cross-scene classification performance.

TABLE II
Class Distribution of Samples for Shanghai–Hangzhou Data Pair

No.	Class Name	Shanghai	Hangzhou
1	Water	123123	18043
2	Land/Building	161689	77450
3	Plant	83188	40207
Total		368000	135700

A. Experimental Data

1) Houston Data Pair: The first cross-scene HSI data pair is Houston data pair, including Houston2013 and Houston2018 datasets. The explicit introduction has been given in the previous section, and their false-color images and ground truth maps are given in Fig. 5, including the one-to-one correspondence between false colors and classes.

2) Shanghai-Hangzhou Data Pair: The second cross-scene HSI data pair is Shanghai–Hangzhou data pair, including Shanghai and Hangzhou dataset [32]. They both were captured by the EO-1 Hyperion hyperspectral sensor [40]. Each of them consists of 198 bands after removing bad bands. Shanghai dataset was captured in April 2002 over Shanghai and consists of 1600×230 pixels. Hangzhou dataset was acquired in November 2002 over Hangzhou and consists of 590×230 pixels. Each contains three classes, including water, land/building, and Plant. The numbers of labeled samples in the two datasets are listed in Table II, and their false-color images and ground truth maps are given in Fig. 6, including the one-to-one correspondence between false colors and classes.

B. Parameter Tuning

SSM is related to two parameters, i.e., neighborhood radius w and iterative times n_t of AEM. We conduct the parameter tuning experiments on the four cases of cross-scene classification, i.e., Houston2013→Houston2018, Houston2018→Houston2018, Shanghai → Hangzhou, and Hangzhou → Shanghai. For all cases, neighborhood radius w is chosen from set $\{1, 2, 3, 4, 5, 6, 7\}$, and the other parameter n_t is also chosen from set $\{1, 2, 3, 4, 5, 6, 7\}$. From Fig. 8, for all cases, the cross-scene performance is acceptable for wide ranges of w and n_t. Compared with Shanghai–Hangzhou datasets, Houston

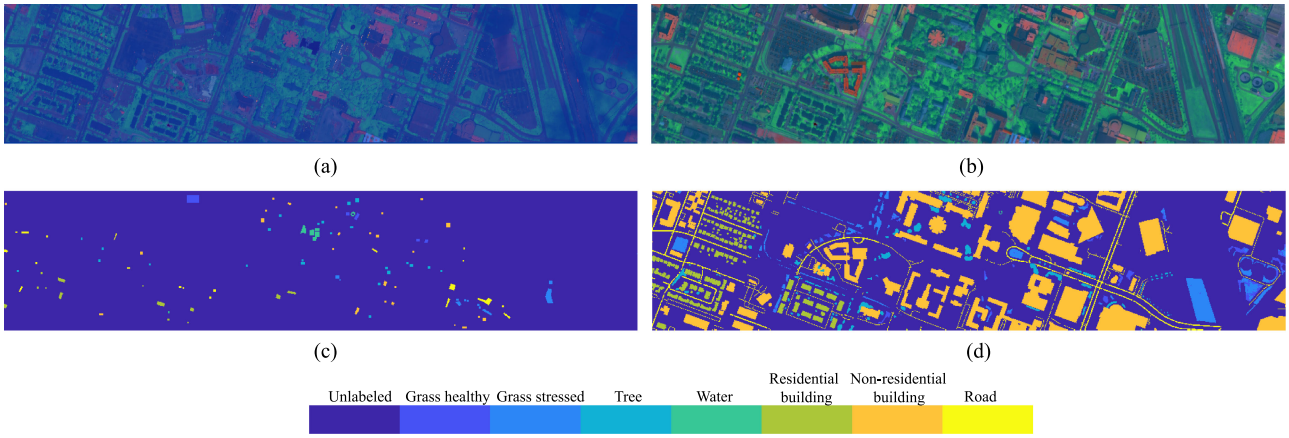

Fig. 5. False-color images and groundtruth maps of Houston data pair. (a) False-color image of Houston2013. (b) False-color image of Houston2018. (c) Groundtruth map of Houston2013. (d) Groundtruth map of Houston2018.

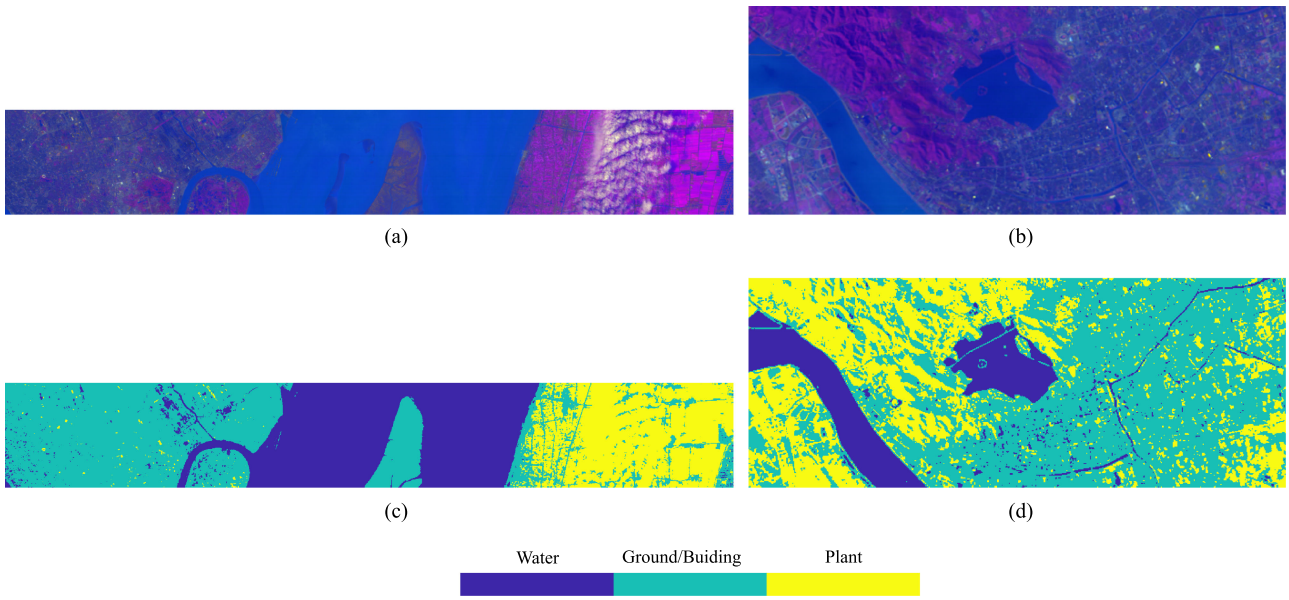

Fig. 6. False-color images and groundtruth maps of Shanghai–Hangzhou data pair. (a) False-color image of Shanghai dataset. (b) False-color image of Hangzhou dataset. (c) Groundtruth map of Shanghai dataset; (d) Groundtruth map of Hangzhou dataset.

datasets are adaptive to relatively large neighborhood radius and more iterative times of AEM. This is because that Houston datasets as urban HSI datasets are more complex, and thus, can benefit a lot from more iterative times of AEM and larger neighborhood radius. After the fine-tuning of parameters $\{w, n_t\}$, the best parameters for four cases are chosen as $\{5, 7\}$, $\{5, 2\}$, $\{1, 2\}$, and $\{1, 2\}$ in subsequent experiments, respectively.

C. Cross-Scene Classification Performance

To validate the effectiveness of the proposed SSM, four experments are conducted in comparison with the aforementioned methods. Four cases, i.e., Houston2013→Houston2018, Houston2018→Houston2018, Shanghai → Hangzhou, and Hangzhou → Shanghai, are listed in Tables III –VI, respectively. The best results are shown in boldface. Since some algorithms like TCA, JDA employ LN as preprocessing, we remain unchanged as described in the original literatures. All methods use SVM as their classifiers.

In all experiments, we can see that cross-scene classification performance of SVM classifier trained directly by the labeled data in source scene is poorer than most other algorithms. This suggests that the discrepancy between source and target scene exists, and it deteriorates cross-scene classification performance. Furthermore, compared with subspace learning algorithms, distribution alignment algorithms are more stable. This can be seen from Tables III–VI that distribution-alignment-based algorithms CORAL, TCA, and JDA averagely outperform JGSA, MEDA, SCA, and STL. This is because methods like JGSA, MEDA, SCA, and STL are more complex and have many hyperparameters to tune, and different hyperparameters result in considerably different classification performance. As

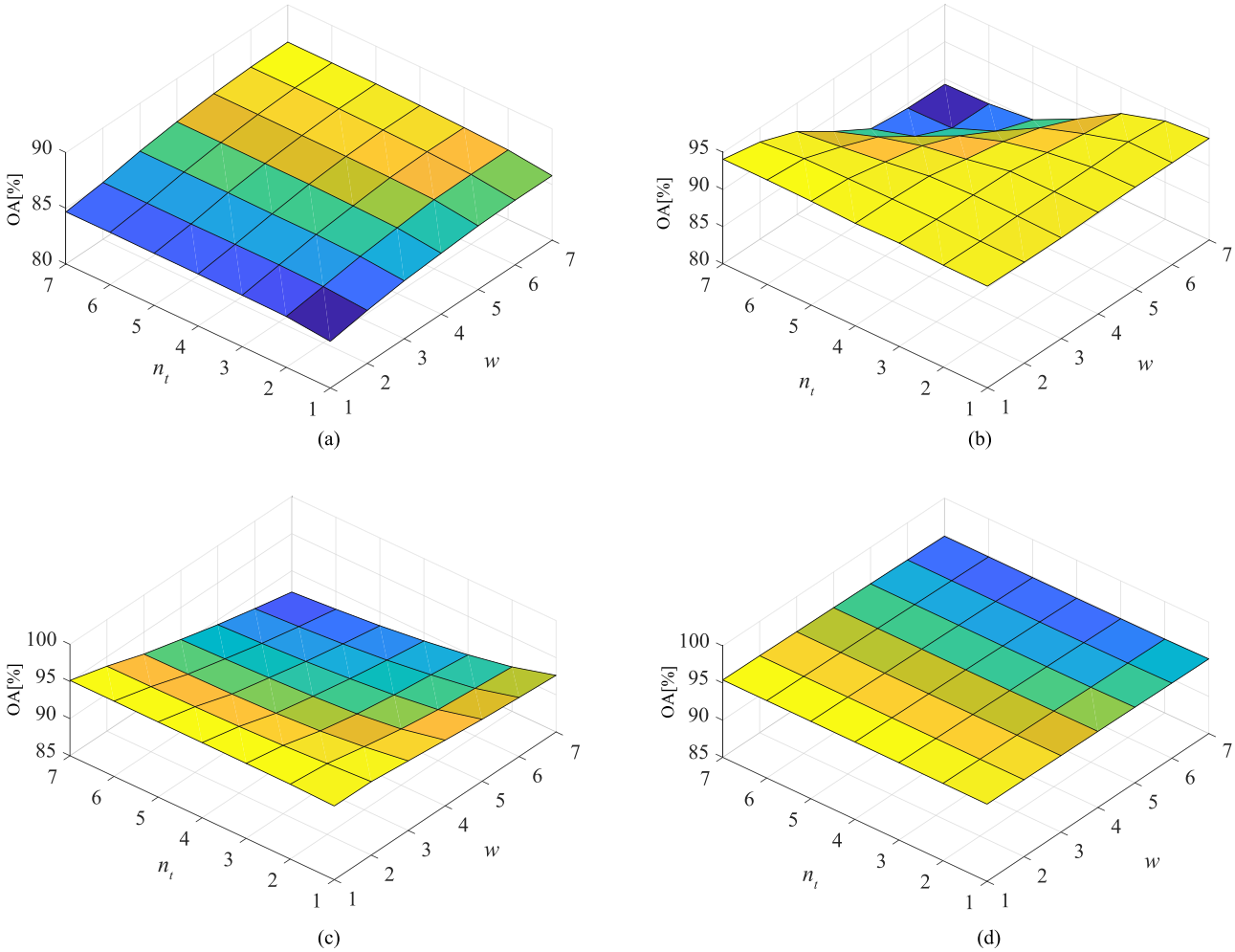

Fig. 7. Parameter tuning of w and n_t for the proposed SSM in four experiments. (a) Houston2013→Houston2018. (b) Houston2018→Houston2013. (c) Shanghai → Hangzhou. (d) Hangzhou → Shanghai.

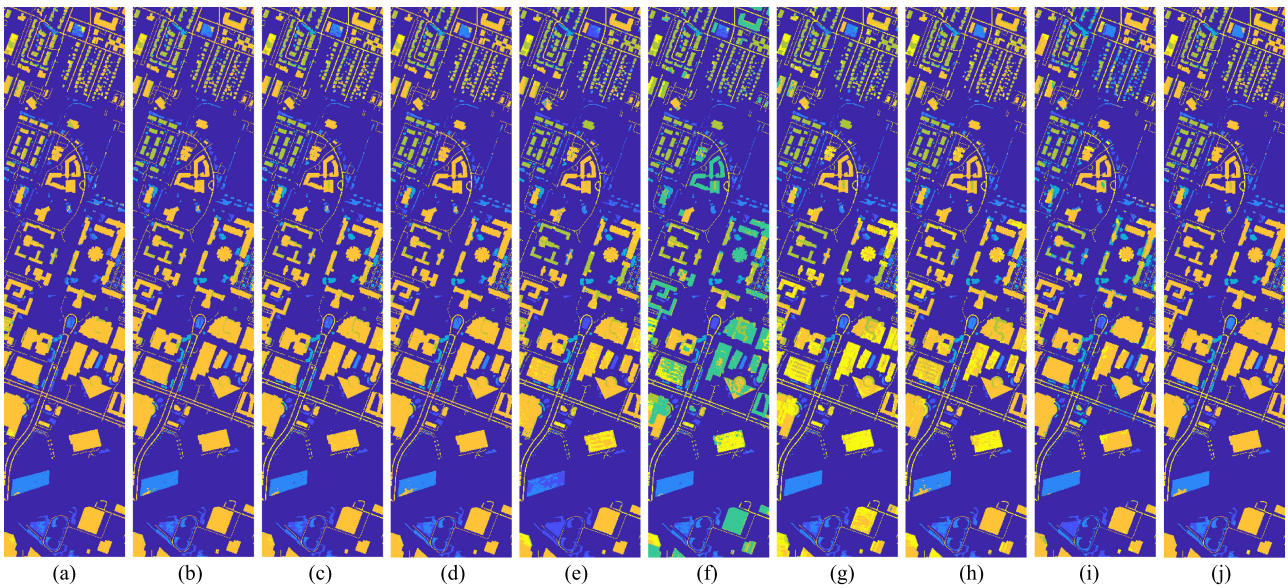

Fig. 8. Cross-scene classification maps by different methods in Houston2013→Houston2018. (a) SVM (72.67%). (b) CORAL (78.27%). (c) TCA (81.76%). (d) JDA (79.24%). (e) JGSA (69.35%). (f) MEDA (38.22%). (g) SCA (53.70%). (h) STL (69.77%). (i) DSAN (70.83%). (j) SSM (86.47%).

TABLE III
CROSS-SCENE CLASSIFICATION PERFORMANCE [%] OF DIFFERENT METHODS ON THE CASE HOUSTON2013→HOUSTON2018

Method	SVM [1]	CORAL [30]	TCA [21]	JDA [24]	JGSA [37]	MEDA [38]	SCA [39]	STL [36]	DSAN [43]	SSM
Class1	98.37	**99.48**	98.82	98.52	99.11	78.86	98.60	86.55	63.78	87.66
Class2	75.43	84.41	78.36	87.87	28.09	76.10	89.30	**92.78**	71.01	92.35
Class3	19.85	57.52	67.17	55.39	**69.70**	63.20	28.81	52.49	74.55	59.83
Class4	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**
Class5	46.81	74.42	74.02	73.97	77.43	**85.60**	79.52	69.70	56.39	**78.75**
Class6	90.37	89.83	86.97	89.88	71.96	13.63	39.45	67.84	76.69	**91.02**
Class7	19.47	22.26	66.91	29.02	74.31	75.02	**78.40**	65.88	52.77	76.51
OA	72.67	78.27	81.76	79.24	69.35	38.22	53.70	69.77	70.83	**86.47**
AA	64.33	75.42	81.75	76.38	74.37	70.34	73.44	76.46	70.74	**83.73**
Kappa	50.77	62.18	70.17	64.00	55.60	31.62	40.20	54.29	55.38	**77.48**

TABLE IV
CROSS-SCENE CLASSIFICATION PERFORMANCE [%] OF DIFFERENT METHODS ON THE CASE HOUSTON2018→HOUSTON2013

Method	SVM [1]	CORAL [30]	TCA [21]	JDA [24]	JGSA [37]	MEDA [38]	SCA [39]	STL [36]	DSAN [43]	SSM
Class1	0.00	95.65	96.52	93.04	84.93	**100**	99.13	98.84	91.30	96.81
Class2	16.99	87.12	92.60	92.88	98.08	55.34	**98.63**	90.41	98.36	97.81
Class3	99.18	97.26	97.26	97.26	94.79	51.23	92.60	62.47	96.44	**98.90**
Class4	90.88	95.79	95.79	95.79	93.68	82.11	93.33	95.79	10.53	**98.60**
Class5	94.98	89.34	91.22	91.54	81.82	75.24	60.82	71.47	65.83	**93.73**
Class6	66.67	64.22	70.34	67.89	73.77	71.32	**99.51**	72.06	98.04	74.51
Class7	97.07	**100**	**100**	**100**	**100**	99.77	0.45	**100**	**100**	**100**
OA	66.72	89.57	91.70	90.91	89.68	76.72	75.42	84.47	83.36	**94.03**
AA	66.54	89.91	91.96	91.20	89.58	76.43	77.78	84.43	80.07	**94.34**
Kappa	61.04	87.78	90.29	89.36	87.91	72.73	71.25	81.80	80.39	**93.01**

TABLE V
CROSS-SCENE CLASSIFICATION PERFORMANCE [%] OF DIFFERENT METHODS ON THE CASE SHANGHAI → HANGZHOU

Method	SVM [1]	CORAL [30]	TCA [21]	JDA [24]	JGSA [37]	MEDA [38]	SCA [39]	STL [36]	DSAN [43]	SSM
Class1	**100**	99.84	99.87	99.67	99.20	97.36	53.74	99.79	87.87	97.11
Class2	86.32	90.05	84.61	85.80	87.70	75.00	72.21	91.52	86.20	**96.31**
Class3	18.68	**93.47**	69.85	71.56	63.17	77.28	78.27	72.91	62.12	92.02
OA	68.10	92.37	82.27	83.43	81.96	78.65	71.55	87.11	79.29	**95.15**
AA	68.33	94.46	84.78	85.68	83.36	83.21	68.07	88.07	78.73	**95.15**
Kappa	40.57	86.91	69.21	71.12	68.22	63.94	52.15	77.22	63.16	**91.44**

TABLE VI
CROSS-SCENE CLASSIFICATION PERFORMANCE [%] OF DIFFERENT METHODS ON THE CASE HANGZHOU → SHANGHAI

Method	SVM [1]	CORAL [30]	TCA [21]	JDA [24]	JGSA [37]	MEDA [38]	SCA [39]	STL [36]	DSAN [43]	SSM
Class1	94.72	95.22	97.55	96.17	98.16	97.15	97.10	96.07	91.90	**98.20**
Class2	59.30	83.36	73.83	70.97	84.08	70.05	77.05	72.30	84.83	**92.75**
Class3	**100**	97.53	**100**	**100**	99.91	97.86	98.42	**100**	95.45	96.43
OA	80.35	90.53	87.68	85.96	92.37	85.40	88.64	86.51	89.60	**95.41**
AA	84.68	92.04	90.46	89.05	94.05	88.35	90.89	89.46	90.73	**95.80**
Kappa	71.17	85.58	81.55	79.04	88.40	78.18	82.88	79.83	84.08	**92.92**

a consequence, the cross-scene classification performance by these methods are unstable. Among distribution-alignment-based methods, CORAL is the most stable since it has no hyperparameter. Deep learning-based method, i.e., DSAN, is only comparative to some subspace learning algorithms. One possible reason for this is that pseudo soft label information in DSAN is calculated by a temporary classifier in each iteration and may be not reliable for target domain, which results in unreliable classification performance. Moreover, the cross-scene classification performance of the proposed SSM is the best in terms of OA, AA, and Kappa in all four cases of cross-scene classification and outperforms the second best one by approximately 4.7% in Houston2013→Houston2018, 2.3% in Houston2018→Houston2013, 2.8% in Shanghai → Hangzhou, and 3% in Hangzhou → Shanghai in terms of OA. In terms of AA and Kappa, the proposed SSM also performs better than all other state-of-the-art methods in all four cases. This shows that good cross-scene classification performance can be achieved by tackling spectral shifts directly, and demonstrates the effectiveness of the proposed SSM.

To visually demonstrate the effectiveness of the proposed SSM, classification maps of all the aforementioned methods are illustrated in Figs. 8 –11. Obviously, the proposed SSM produces the most accurate and the spatially smoothest classification maps with less mislabeled pixels, which are consistent with the results listed in Tables III–VI.

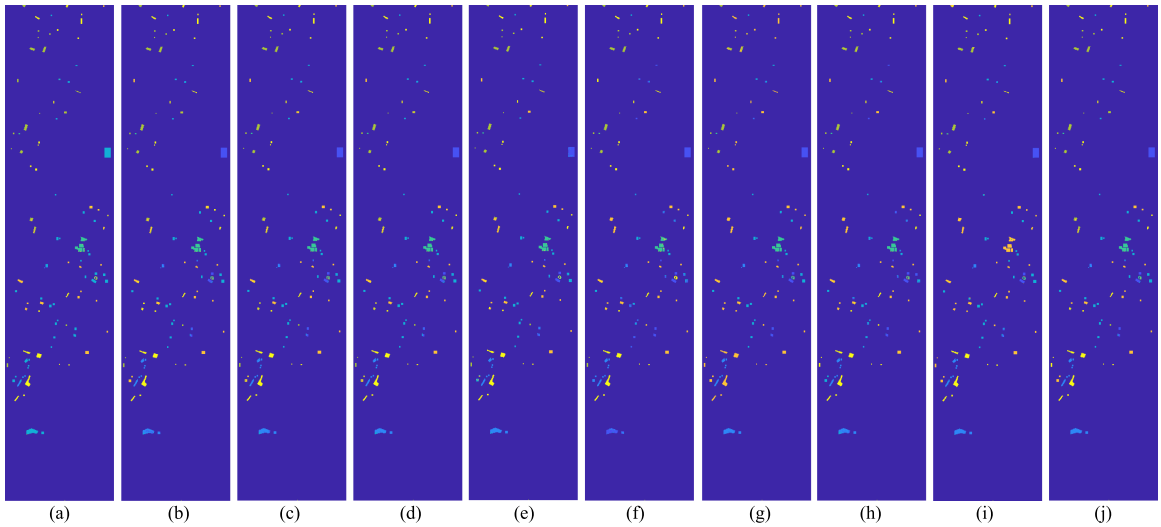

Fig. 9. Cross-scene classification maps by different methods in Houston2018→Houston2013. (a) SVM (66.72%). (b) CORAL (89.57%). (c) TCA (91.70%). (d) JDA (90.91%). (e) JGSA (89.68%). (f) MEDA (76.72%). (g) SCA (75.42%). (h) STL (84.47%). (i) DSAN (83.36%). (j) SSM (94.03%).

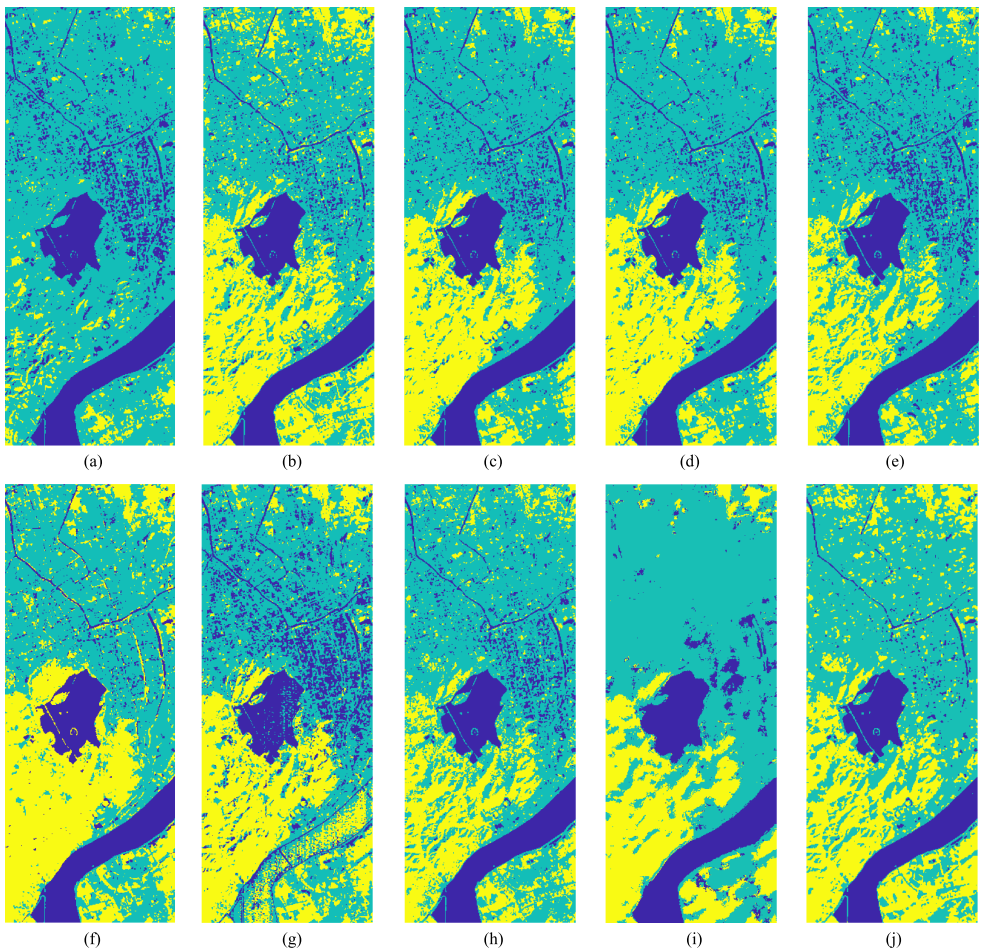

Fig. 10. Cross-scene classification maps by different methods in Shanghai → Hangzhou. (a) SVM (68.10%). (b) CORAL (92.37%). (c) TCA (82.27%). (d) JDA (83.43%). (e) JGSA (81.96%). (f) MEDA (78.65%). (g) SCA (71.55%). (h) STL (87.11%). (i) DSAN (79.29%). (j) SSM (95.15%).

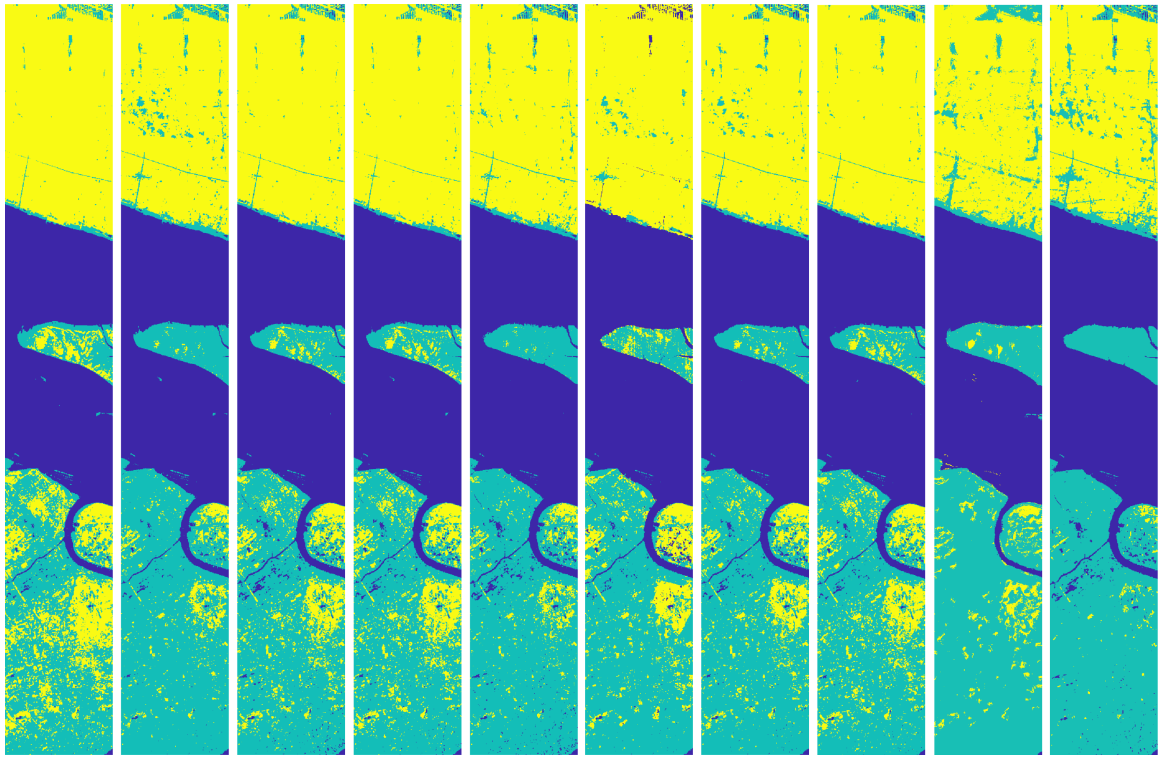

Fig. 11. Cross-scene classification maps by different methods in Hangzhou → Shanghai. (a) SVM (80.35%). (b) CORAL (90.53%). (c) TCA (87.68%). (d) JDA (85.96%). (e) JGSA (92.37%). (f) MEDA (85.40%). (g) SCA (88.64%). (h) STL (86.51%). (i) DSAN (89.60%). (j) SSM (95.41%).

TABLE VII
COMPUTATION TIME (IN SECONDS) OBTAINED THROUGH DIFFERENT METHODS IN HOUSTON2013→HOUSTON2018

Method	Computation Time(s)
CORAL	1.2
TCA	6
JDA	18
JGSA	80.5
MEDA	450
SCA	50.3
STL	11.5
DSAN	21.0
SSM	13.5

To illustrate the computational complexity of the proposed SSM compared to other methods for cross-scene HSI classification, Table VII lists the computation time of several methods in Houston2013→Houston2018 by using MATLAB on an Intel(R) Core(TM) i5-7300HQ central processing unit with 8 GB of RAM. For the proposed SSM, the computation time is the execution time of AN and AEM with neighborhood radius of 5 for one time. As shown in Table VII, among all the methods for cross HSI classification, MEDA costs the longest time, followed by JGSA, which costs 80.5 s. The computation time of the remaining methods is below 60 s. CORAL takes only 1.2 s. Deep learning-based method DSAN costs 21 s which is the test time, but its training time takes hours. The proposed SSM costs only 13.5 s, which shows the efficiency of the proposed SSM. The computation time of the proposed SSM mainly lies on the execution of AEM. For less neighborhood radius, it costs less time.

V. CONCLUSION

In this article, a simple preprocessing method called SSM has been proposed for cross-scene HSI classification, including ASM and AEM. ASM is to reduce the spectral shift caused by amplitude shift, i.e., the shape of spectral curve of one material is stable while the amplitude varies considerably. AEM is to reduce the spectral shift in target scene caused by adjacency effect, i.e., the spectral curve of a pixel of one class is affected by nearby pixels of different classes in the scene. Experiments on four cases generated by two cross-scene HSI data pairs have shown that the proposed SSM outperforms the existing state-of-the-art methods for cross-scene classification, and thus its effectiveness has been demonstrated.

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] M. Zhang, W. Li, and Q. Du, "Diverse region-based CNN for hyperspectral image classification," *IEEE Trans. Image Process.*, vol. 27, no. 6, pp. 2623–2634, Jun. 2018.
[3] W. Li, E. W. Tramel, S. Prasad, and J. E. Fowler, "Nearest regularized subspace for hyperspectral classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 1, pp. 477–489, Jan. 2014.
[4] N. H. Ly, Q. Du, and J. E. Fowler, "Collaborative graph-based discriminant analysis for hyperspectral imagery," *IEEE J. Sel, Top. Appl. Earth Observ. Remote Sens.*, vol. 7, no. 6, pp. 2688–2696, Jun. 2014.

[5] W. Li and Q. Du, "Laplacian regularized collaborative graph for discriminant analysis of hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 12, pp. 7066–7076, Dec. 2016.

[6] K. Makantasis, A. D. Doulamis, N. D. Doulamis, and A. Nikitakis, "Tensor-based classification models for hyperspectral data analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 12, pp. 6884–6898, Dec. 2018.

[7] K. Makantasis, K. Karantzalos, A. Doulamis, and N. Doulamis, "Deep supervised learning for hyperspectral data classification through convolutional neural networks," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2015, pp. 4959–4962.

[8] Y. Chen, Z. Lin, X. Zhao, G. Wang, and Y. Gu, "Deep learning-based classification of hyperspectral data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 6, pp. 2094–2107, Jun. 2014.

[9] X. Yang, Y. Ye, X. Li, R. Y. Lau, X. Zhang, and X. Huang, "Hyperspectral image classification with deep learning models," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 9, pp. 5408–5423, Sep. 2018.

[10] D. Hong, L. Gao, J. Yao, B. Zhang, A. Plaza, and J. Chanussot, "Graph convolutional networks for hyperspectral image classification," *IEEE Trans Geosci. Remote Sens.*, 2020, doi: 10.1109/TGRS.2020.3015157.

[11] D. Hong, N. Yokoya, J. Chanussot, and X. X. Zhu, "An augmented linear mixing model to address spectral variability for hyperspectral unmixing," *IEEE Trans. Image Process.*, vol. 28, no. 4, pp. 1923–1938, Apr. 2019.

[12] D. Hong, X. Wu, P. Ghamisi, J. Chanussot, N. Yokoya, and X. X. Zhu, "Invariant attribute profiles: A spatial-frequency joint feature extractor for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 6, pp. 3791–3808, Jun. 2020.

[13] D. Hong, N. Yokoya, J. Chanussot, J. Xu, and X. X. Zhu, "Learning to propagate labels on graphs: An iterative multitask regression framework for semi-supervised hyperspectral dimensionality reduction," *ISPRS J. Photogrammetry Remote Sens.*, vol. 158, pp. 35–49, 2019.

[14] J. R. Schott, C. Salvaggio, and W. J. Volchok, "Radiometric scene normalization using pseudoinvariant features," *Remote Sens. Environ.*, vol. 26, no. 1, pp. 1–16, 1988.

[15] Z. Lee, K. L. Carder, S. K. Hawes, R. G. Steward, T. G. Peacock, and C. O. Davis, "Model for the interpretation of hyperspectral remote-sensing reflectance," *Appl. Opt.*, vol. 33, no. 24, pp. 5721–5732, 1994.

[16] E. Vermote *et al.*, "Atmospheric correction of visible to middle-infrared EOS-MODIS data over land surfaces: Background, operational algorithm and validation," *J. Geophysical Res.: Atmospheres*, vol. 102, no. D14, pp. 17131–17141, 1997.

[17] E. F. Vermote, D. Tanré, J. L. Deuze, M. Herman, and J.-J. Morcette, "Second simulation of the satellite signal in the solar spectrum, 6S: An overview," *IEEE Trans. Geosci. Remote Sens.*, vol. 35, no. 3, pp. 675–686, May 1997.

[18] G. A. Shaw and H. K. Burke, "Spectral imaging for remote sensing," *Lincoln Lab. J.*, vol. 14, no. 1, pp. 3–28, 2003.

[19] J. G. Masek *et al.*, "A Landsat surface reflectance dataset for North America, 1990-2000," *IEEE Geosci. Remote Sens. Lett.*, vol. 3, no. 1, pp. 68–72, Jan. 2006.

[20] N. E. Young, R. S. Anderson, S. M. Chignell, A. G. Vorster, R. Lawrence, and P. H. Evangelista, "A survival guide to Landsat preprocessing," *Ecology*, vol. 98, no. 4, pp. 920–932, 2017.

[21] S. J. Pan, I. W. Tsang, J. T. Kwok, and Q. Yang, "Domain adaptation via transfer component analysis," *IEEE Trans. Neural Netw.*, vol. 22, no. 2, pp. 199–210, Feb. 2011.

[22] B. Geng, D. Tao, and C. Xu, "DAML: Domain adaptation metric learning," *IEEE Trans. Image Process.*, vol. 20, no. 10, pp. 2980–2989, Oct. 2011.

[23] M. Baktashmotlagh, M. T. Harandi, B. C. Lovell, and M. Salzmann, "Unsupervised domain adaptation by domain invariant projection," in *Proc. IEEE Int. Conf. Comput. Vision*, 2013, pp. 769–776.

[24] M. Long, J. Wang, G. Ding, J. Sun, and P. S. Yu, "Transfer feature learning with joint distribution adaptation," in *Proc. IEEE Int. Conf. Comput. Vision*, 2013, pp. 2200–2207.

[25] Z. Sun, C. Wang, H. Wang, and J. Li, "Learn multiple-kernel SVMs for domain adaptation in hyperspectral data," *IEEE Geosci. Remote Sens. Lett.*, vol. 10, no. 5, pp. 1224–1228, Sep. 2013.

[26] G. Matasci, M. Volpi, M. Kanevski, L. Bruzzone, and D. Tuia, "Semisupervised transfer component analysis for domain adaptation in remote sensing image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 7, pp. 3550–3564, Jul. 2015.

[27] J. Xia, N. Yokoya, and A. Iwasaki, "Ensemble of transfer component analysis for domain adaptation in hyperspectral remote sensing image classification," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2017, pp. 4762–4765.

[28] J. Peng, W. Sun, L. Ma, and Q. Du, "Discriminative transfer joint matching for domain adaptation in hyperspectral image classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 6, pp. 972–976, Jun. 2019.

[29] M. Long, J. Wang, G. Ding, J. Sun, and P. S. Yu, "Transfer joint matching for unsupervised domain adaptation," in *Proc. IEEE Conf. Comput. Vision Pattern Recognit.*, 2014, pp. 1410–1417.

[30] B. Sun and K. Saenko, "Deep CORAL: Correlation alignment for deep domain adaptation," in *Proc. Eur. Conf. Comput. Vision*, 2016, pp. 443–450.

[31] L. Ma, M. M. Crawford, L. Zhu, and Y. Liu, "Centroid and covariance alignment-based domain adaptation for unsupervised classification of remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 4, pp. 2305–2323, Apr. 2019.

[32] M. Ye, Y. Qian, J. Zhou, and Y. Y. Tang, "Dictionary learning-based feature-level domain adaptation for cross-scene hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 3, pp. 1544–1562, Mar. 2017.

[33] A. Samat, P. Gamba, J. Abuduwaili, S. Liu, and Z. Miao, "Geodesic flow kernel support vector machine for hyperspectral image classification by unsupervised subspace feature transfer," *Remote Sens.*, vol. 8, no. 3, 2016, Art. no. 234.

[34] B. Gong, Y. Shi, F. Sha, and K. Grauman, "Geodesic flow kernel for unsupervised domain adaptation," in *Proc. IEEE Conf. Comput. Vision Pattern Recognit.*, 2012, pp. 2066–2073.

[35] T. Liu, X. Zhang, and Y. Gu, "Unsupervised cross-temporal classification of hyperspectral images with multiple geodesic flow kernel learning," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 12, pp. 9688–9701, Dec. 2019.

[36] J. Wang, Y. Chen, L. Hu, X. Peng, and S. Y. Philip, "Stratified transfer learning for cross-domain activity recognition," in *Proc. IEEE Int. Conf. Pervasive Comput. Commun.*, 2018.

[37] J. Zhang, W. Li, and P. Ogunbona, "Joint geometrical and statistical alignment for visual domain adaptation," in *Proc. IEEE Conf. Comput. Vision Pattern Recognit.*, 2017, pp. 1859–1867.

[38] J. Wang, W. Feng, Y. Chen, H. Yu, M. Huang, and P. S. Yu, "Visual domain adaptation with manifold embedded distribution alignment," in *Proc. 26th ACM Int. Conf. Multimedia*, 2018, pp. 402–410.

[39] M. Ghifary, D. Balduzzi, W. B. Kleijn, and M. Zhang, "Scatter component analysis: A unified framework for domain adaptation and domain generalization," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 7, pp. 1414–1430, Jul. 2017.

[40] D. Tuia, E. Pasolli, and W. J. Emery, "Using active learning to adapt remote sensing image classifiers," *Remote Sens. Environ.*, vol. 115, no. 9, pp. 2232–2242, 2011.

[41] C. Persello and L. Bruzzone, "Active learning for domain adaptation in the supervised classification of remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 11, pp. 4468–4483, Nov. 2012.

[42] M. Long, Y. Cao, J. Wang, and M. Jordan, "Learning transferable features with deep adaptation networks," in *Proc. Int. Conf. Mach. Learn.*, 2015, pp. 97–105.

[43] Y. Zhu *et al.*, "Deep subdomain adaptation network for image classification," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 4, pp. 1713–1722, Apr. 2021.

[44] C. Debes *et al.*, "Hyperspectral and LiDAR data fusion: Outcome of the 2013 GRSS data fusion contest," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 6, pp. 2405–2418, Jun. 2014.

[45] Y. Xu *et al.*, "Advanced multi-sensor optical remote sensing for urban land use and land cover classification: Outcome of the 2018 IEEE GRSS data fusion contest," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 6, pp. 1709–1724, Jun. 2019.

[46] G. Schaepman-Strub, M. E. Schaepman, T. H. Painter, S. Dangel, and J. V. Martonchik, "Reflectance quantities in optical remote sensing-definitions and case studies," *Remote Sens. Environ.*, vol. 103, no. 1, pp. 27–42, 2006.

[47] D. Schläpfer and R. Richter, "Evaluation of BREFCOR BRDF effects correction for HYSPEX, CASI, and APEX imaging spectroscopy data," in *Proc. 6th Workshop Hyperspectral Image Signal Process.: Evol. Remote Sens.*, 2014.

[48] R. Richter and D. Schläpfer, "Geo-atmospheric processing of airborne imaging spectrometry data. Part 2: Atmospheric/topographic correction," *Int. J. Remote Sens.*, vol. 23, no. 13, pp. 2631–2649, 2002.

[49] A. Berk, L. S. Bernstein, and D. C. Robertson, "MODTRAN: A moderate resolution model for LOWTRAN," Spectral Science Inc., Burlington, MA, USA, Tech. Rep. SSI-TR-124, 1987.

Huan Liu received the B.S. degree from China University of Mining and Technology, Xuzhou, China, and the M.S. degree from Beijing Institute of Technology, Beijing, China, in 2013 and 2016, respectively. He is working toward the Ph.D. degree in the School of Information and Electronics, Beijing Institute of Technology, Beijing, China. His research interests include image processing, image classification, and transfer learning.

Wei Li (Senior Member, IEEE) received the B.E. degree in telecommunications engineering from Xidian University, Xi'an, China, in 2007, the M.S. degree in information science and technology from Sun Yat-sen University, Guangzhou, China, in 2009, and the Ph.D. degree in electrical and computer engineering from Mississippi State University, Starkville, MS, USA, in 2012.

Subsequently, he spent one year as a Postdoctoral Researcher with the University of California at Davis, Davis, CA, USA. He is a Professor with the School of Information and Electronics, Beijing Institute of Technology, Beijing, China. His research interests include hyperspectral image (HSI) analysis, pattern recognition, and data compression.

Dr. Li was the recipient of the 2015 Best Reviewer Award from the IEEE Geoscience and Remote Sensing Society (GRSS) for his service for the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING. He is serving as an Associate Editor for the IEEE SIGNAL PROCESSING LETTERS. He served as a Guest Editor for special issue of the *Journal of Real-Time Image Processing*, *Remote Sensing*, and the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING. He is currently an Associate Editor for the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, and the IEEE SIGNAL PROCESSING LETTERS.

Xiang-Gen Xia (Fellow, IEEE) received the B.S. degree in mathematics from Nanjing Normal University, Nanjing, China, and the M.S. degree in mathematics from Nankai University, Tianjin, China, and the Ph.D. degree in electrical engineering from the University of Southern California, Los Angeles, in 1983, 1986, and 1992, respectively.

He was a Senior/Research Staff Member with Hughes Research Laboratories, Malibu, CA, USA, during 1995–1996. In September 1996, he joined the Department of Electrical and Computer Engineering, University of Delaware, Newark,, DE, USA, where he is the Charles Black Evans Professor. Dr. Xia has authored the book Modulated Coding for Intersymbol Interference Channels (Marcel Dekker, 2000). His research interests include space-time coding, MIMO and OFDM systems, digital signal processing, and SAR and ISAR imaging.

Dr. Xia was the recipient of the National Science Foundation (NSF) Faculty Early Career Development (CAREER) Program Award, in 1997, the Office of Naval Research (ONR) Young Investigator Award, in 1998, and the Outstanding Overseas Young Investigator Award from the National Nature Science Foundation of China, in 2001. He was the recipient of the 2019 Information Theory Outstanding Overseas Chinese Scientist Award, The Information Theory Society of Chinese Institute of Electronics. Dr. Xia has served as an Associate Editor for numerous international journals including IEEE TRANSACTIONS ON SIGNAL PROCESSING, IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, IEEE TRANSACTIONS ON MOBILE COMPUTING, and IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY. Dr. Xia is Technical Program Chair of the Signal Processing Symposium, IEEE Global Communications Conference (Globecom) 2007 in Washington D.C., and the General Co-Chair of International Conference on Acoustics, Speech, & Signal Processing (ICASSP) 2005 in Philadelphia.

Mengmeng Zhang received the B.S. degree in computer science and technology from the Qingdao University of Science and Technology, Qingdao, China, in 2014, and the Ph.D. degree in control science and engineering from Beijing University of Chemical Technology, Beijing, China, in 2019.

She is currently a Postdoctoral Researcher of the School of Information and Electronics with Beijing Institute of Technology, Beijing, China. Her research interests include remote sensing image processing and pattern recognition, remote sensing image process, and pattern recognition.

Chen-Zhong Gao received the B.S. degree in information and communication engineering from Beijing Institute of Technology, Beijing, China, where he is currently working toward the Ph.D. degree with the School of Information and Electronics.

His research interests include remote sensing and image processing.

Ran Tao (Senior Member, IEEE) was born in 1964. He received the B.S. degree from the Electronic Engineering Institute of PLA, Hefei, China, in 1985, and the M.S. and Ph.D. degrees from the Harbin Institute of Technology, Harbin, China, in 1990 and 1993, respectively.

In 2001, he was a Senior Visiting Scholar with the University of Michigan, Ann Arbor, MI, USA. He is a Professor with the School of Information and Electronics, Beijing Institute of Technology, Beijing, China. He was a Distinguished Professor of the Changjiang Scholars Program in 2009. He was a Chief Professor of the Program for Changjiang Scholars and Innovative Research Team in University from 2010 to 2012. He has been a Chief Professor of the Creative Research Groups of the National Natural Science Foundation of China since 2014. He has authored or coauthored three books and over 100 peer-reviewed journal articles. His research interests include fractional Fourier transform and its applications, theory, and technology for radar and communication systems.

Dr. Tao is a Member of the Wireless Communication and Signal Processing Commission of International Union of Radio Science (URSI). He was the recipient of the National Science Foundation of China for Distinguished Young Scholars in 2006, and the First Prize of Science and Technology Progress in 2006 and 2007, and the First Prize of Natural Science in 2013, both awarded by the Ministry of Education. He is an Associate Editor for the IEEE SIGNAL PROCESSING LETTERS. He is the Vice Chair of the IEEE China Council and the URSI China Council.