SPATIAL ENTROPY BASED MUTUAL INFORMATION IN HYPERSPECTRAL BAND SELECTION FOR SUPERVISED CLASSIFICATION

BAIJIE WANG, XIN WANG, AND ZHANGXIN CHEN

Abstract. Hyperspectral band image selection is a fundamental problem for hyperspectral remote sensing data processing. Accepting its importance, several information-based band selection methods have been proposed, which apply Shannon entropy to measure image information. However, the Shannon entropy is not accurate in measuring image information since it neglects the spatial distribution of pixels and is computed only from a histogram. This paper investigates the potential of spatial entropy in measuring image information and proposes a new mutual information (MI) band selection method based on the spatial entropy. Then selected band images are validated for supervised classification via Support Vector Machine (SVM). Using a hyperspectral AVIRIS 92AV3C dataset, experiment results show that with 20 images selection from 220 bands, the supervised classification accuracy can reach 90.6%. Comparison with a previous Shannon entropy-based band selection method shows that the proposed method selects band images which can achieve more accurate classification results.

Key words. Spatial entropy, mutual information, band selection, support vector machine, classification, hyperspectral remote sensing data

1. Introduction

Hyperspectral sensors measure hundreds of contiguous spectral bands with narrow spectral intervals simultaneously, which can provide fine, detailed and large volume spectral information. For example, an Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) is a premier instrument in the realm of Earth Remote Sensing [9]. It is a unique optical sensor that delivers calibrated images in 224 contiguous spectral bands with wavelengths from 0.4 μm to 2.5 μm. Hyperspectral sensors benefit the potential to detect targets and classify materials with high accuracy.

High dimensionality of the hyperspectral remote sensing images also calls for effective and efficient feature selection methods. For example, for a land use classification task, it is unnecessary to process all spectral bands from the hyperspectral images since some bands may contain less discriminatory information than the others. Besides, the computational cost for hyperspectral image processing with all bands is high; e.g., as the dimension increases, say 224, the computational cost for classification will be unendurable. Therefore, it is an advantage to identify bands that conveys more information.

Band selection refers to the selection of band images with relevant information or with weak correlation [6]. Information-based band selection is an active research topic recently [3, 5, 7, 10], which generally applies Shannon entropy or its variations, e.g., mutual information (MI), as the measurement evaluating the image information. In [10], the Shannon entropy was directly applied to band selection. Band images are ranked in terms of entropy values and ones with higher entropy

values are selected. Since the Shannon entropy is computed based on a single image, without reference to any ground truth or benchmark, its variation, mutual information (MI), is introduced. MI measures the information shared between each band image and a reference, i.e., the ground truth, and images with higher MI values are selected. In [3], the authors introduced a method using MI based clustering to deal with multispectral images selection. In [7], a MI estimation method was introduced and the band selection method was developed with the objective to choose band images which maximizes the joint MI value.

Each pixel in remote sensing images has spatial attributes, such as row and column, and non-spatial attributes, such as intensity. Even though the existing information-based band selection methods based on Shannon entropy can give good results in many cases, the Shannon entropy is calculated only based on the statistics of the non-spatial attributes. Therefore, it leads to an incomplete evaluation of image information by ignoring spatial distribution of pixels. We will illustrate the consequent problem by the following example.
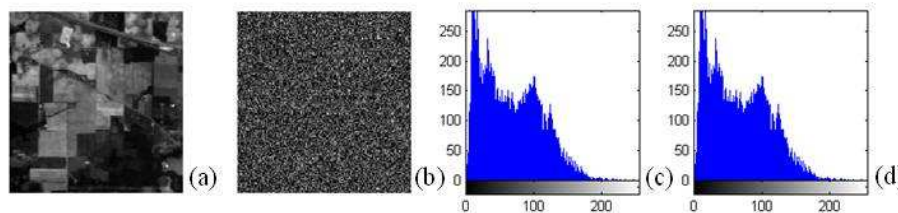

FIGURE 1. Comparison between a qualified remote sensing image and a noise image with the same histogram: (a) Qualified image, (b) Noise image, (c) Histogram of (a), (d) Histogram of (b).

Fig.1 (a) shows a remote sensing image of an area and Fig.1 (b) is a noise image of the same area. Given an image classification task, Fig.1 (b) gives no valuable information since all pixels are noise contaminated while Fig.1 (a) may give relevant information since distinct patterns can be visually observed. Figs.1 (c) and (d) are the image histograms of Figs.1 (a) and (b), respectively. The two histograms indicate that two images have the identical intensity distribution.

In this example, the two images have the identical histograms; if we use the Shannon entropy to measure the image information in Figs.1 (a) and (b), their entropy would be the same value of 0.4481. In this case, the Shannon entropy fails in discriminating the information difference between Figs.1 (a) and (b).

Spatial entropy is the extension of the Shannon entropy with the spatial configuration, which measures the distribution of a non-spatial attribute in the spatial domain [4, 14]. In this paper, we propose to use spatial entropy measuring band image information and develop a new information-based band selection method considering both the pixels' intensity and the spatial location in an image. The main contributions are summarized below:

First, this paper introduces a new spatial entropy-based mutual information (SEMI) function as the image information measurement. The SEMI function is derived from the spatial entropy model in [4, 14] and the classic mutual information definition [11]. SEMI quantifies the shared information between the band image and the reference image with the images' spatial and non-spatial attributes, i.e., the pixels' intensity and location.

Second, this paper develops a new band selection method. This new band selection method selects relevant band images from candidates with high SEMI values and low image correlation.

Finally, the proposed band selection method is demonstrated with supervised classification tasks. A hyperspectral AVIRIS 92AV3C remote sensing dataset, containing 220 band images and 1 ground truth image, is chosen for the experiment. Selected band images are used as source data for the supervised land use classification with the current popular support vector machine (SVM) classifier.

The remainder of the paper is organized as follows. Section 2 introduces a spatial entropy-based mutual information (SEMI) function and the new band selection method. Section 3 presents the supervised support vector machine (SVM) classification experiments using the band selection results from the AVIRIS 92AV3C dataset. Finally, Section 4 draws conclusions and discusses our future work.

2. Methodology

In this section, we propose a new spatial entropy-based band selection method. Specifically, Sections 2.1 and 2.2 introduce the spatial entropy-based mutual information (SEMI) as the band selection criterion; Section 2.3 introduces a new band selection algorithm. In Section 2.4, we review the support vector machine for supervised land use classification.

Let $D = \{d_i | i = 1, 2, \cdots, n\}$ be n band images acquired from a hyperspectral remote sensor. Each image $d_i \in D$ can be viewed as a spatial dataset where each pixel of the image has spatial attributes, i.e., the row and column, and a non-spatial attribute, i.e., the intensity value. The objective of the band selection is to identify a subset SD from D such that $SD = \{d_j | j = 1, 2, \cdots, m\}, m \leq n$, contains relevant features for the land use classification.

2.1. Shannon Entropy and Mutual Information. Before introducing spatial entropy-based mutual information (SEMI), it is necessary to review Shannon entropy and the classic mutual information.

Given an image X containing N pixels, each pixel has its intensity value G and the location attributes *row* and *col*. First, a partition process assigns each pixel in X to a unique slot based on its intensity value G, which generates a partition of X, denoted by $(X_1, \cdots, X_i, \cdots, X_t)$. For example, given an image with the radiometric resolution of eight bits, the intensity G of pixels will have 256 levels so that all the pixels in the image can be binned into 256 slots. Let $(p_1, \cdots, p_i, \cdots, p_t)$ be the fraction of the number of pixels in category X_i over N, i.e., $p_i = |Xi|/N$ and $\sum p_i = 1$. In practice, the computation of $(p_1, \cdots, p_i, \cdots, p_t)$ is an approximated by the image histogram.

Definition 1: Given the partition $(X_1, \cdots, X_i, \cdots, X_t)$ of an image X, the Shannon entropy of X is defined as:

$$(1) \qquad H(X) = -\sum_{i=1}^{t} p_i log_2(p_i)$$

where p_i is the fraction of the number of pixels in category X_i over N. In Eq. (1), the Shannon entropy is computed from the image histogram and so a statistical information measurement of image X.

Shannon entropy H is defined for a single band image and it is not related to the target information. For example, in a land use classification task, apart from the remote sensing images, partial ground truth can also be acquired as target

information, or reference data. Since the Shannon entropy does not involve the target information, it is seldom used to evaluate the relevance of a band image to the target land use classification task. Hence the variation of the Shannon entropy, mutual information (MI), is introduced to measure the relevance of a single band image to the target image.

Definition 2: Given the band image X and the ground truth image Y, the mutual information of these two images is defined as:

$$(2) \qquad I(X,Y) = H(X) + H(Y) - H(X,Y)$$

In Definition 2, $H(X)$ and $H(Y)$ are the Shannon entropies to the band image X and the ground truth image Y, respectively, and $H(X,Y)$ is the joint entropy between X and Y. MI measures the common information shared by X and Y. The higher MI value X and Y share, the more relevant X is to the land use classification. Therefore, the high value of MI is a general information-based band selection criterion.

However, as in Eq. (2), MI is defined with the Shannon entropy, which ignores the spatial information inside the image. For example, MI is computed from the image histograms of X and Y, and the pixels' spatial attribute is discarded. Therefore, either the Shannon entropy or MI evaluation is an incomplete image information measurement by ignoring spatial attributes inside a band image.

2.2. Spatial Entropy-based Mutual Information. Spatial entropy is an extension of Shannon entropy which has a spatial configuration. Various forms of the spatial entropy have been developed and this paper selects the one from [4] because it is simple to compute. In the spatial entropy, the spatial configuration is defined as the ratio of the average distance between objects of a particular class (named the intra-distance) to the average distance of a particular class to the others (named the extra-distance). The intra-distance of X_i, denoted by d_i^{int} is the average distance between pixels in X_i (as shown in Eq. (3)). The extra-distance of X_i, denoted by d_i^{ext} is the average distance of pixels in X_i to other partition classes of X, as shown in Eq. (4).

$$(3) \qquad d_i^{int} = \begin{cases} \frac{1}{|X_i| \times |X_i - 1|} \sum_{j=1, j \in X_i}^{|X_i|} \sum_{k=1, k \neq j, k \in X_i}^{|X_i|} dist(j,k) & \text{if} |X_i| > 1 \\ \lambda & \text{otherwise} \end{cases}$$

$$(4) \qquad d_i^{ext} = \begin{cases} \frac{1}{|X_i| \times |X - X_i|} \sum_{j=1, j \in X_i}^{|X_i|} \sum_{k=1, k \neq j, k \notin X_i}^{|X - X_i|} dist(j,k) & \text{if} X \neq X_i \\ \beta & \text{otherwise} \end{cases}$$

In Eq. (3), when X_i is empty or contains only one object, we assume its intra-distance is very small and a small constant λ is assigned to d_i^{int} to avoid the influence of null values on the computation. In Eq. (4), when X_i includes all of the objects in X, i.e., all objects have the same values of G, we assume that the extra-distance d_i^{ext} is very large, and assign the extra-distance with a large constant β. $dist(j,k)$ is the distance between objects j and k in the spatial spaces.

Definition 3: The spatial entropy of an image X based on its partition $\{X_1, \cdots, X_i, \cdots, X_t\}$ is defined as:

(5)
$$H_s(X) = -\sum_{i=1}^{t} \frac{d_i^{int}}{d_i^{ext}} p_i \log_2(p_i)$$

In this definition, a spatial configuration d_i^{int}/d_i^{ext} is added as a weight factor in the Shannon entropy, which enables the spatial entropy H_s to measure the image information by evaluating the pixels' non-spatial attribute in a spatial space.

The Shannon entropy, as in Eq. (1), measures the probability distribution of the pixels' non-spatial attribute G, i.e., the intensity value. The Spatial entropy H_s is a special form of the Shannon entropy and has a spatial configuration d_i^{int}/d_i^{ext}. Even though each pixel's intensity value is correlated within the spatial spaces, the probability distribution of G is independent of d_i^{int}/d_i^{ext}. Hence, the weight factor d_i^{int}/d_i^{ext} does not influence the property of H_s in measuring the probability distribution of G.

On the other hand, the spatial entropy H_s takes the spatial correlation into consideration while measuring the image information. A spatial correlation generally exists among a spatial dataset. The First Law of Geography [12] states "everything is related to everything else, but near things are more related than distant things". As a spatial dataset, the spatial attributes and the spatial correlation in remote sensing images cannot be neglected. This is also the reason why the Shannon entropy fails in discriminating the information difference between the qualified and the noise image in Fig.1. The spatial entropy has the spatial configuration d_i^{int}/d_i^{ext} that keeps the spatial entropy decreasing when pixels with similar intensity values are close and diverse pixels are far from each other. For a band image of which similar pixels are close and dissimilar pixels are far from each other, d_i^{int}/d_i^{ext} decreases. Therefore, the spatial entropy is a better image information measurement by incorporating both spatial and non-spatial attributes inside the band image.

Spatial entropy itself cannot provide the relevance from the band image to the reference image, so we introduce the spatial entropy-based mutual information (SEMI) below. Substituting H with H_s in Eq. (2), the SEMI is defined:

Definition 4: Given the band image X and the ground truth image Y, the Spatial entropy-based mutual information equals:

(6)
$$I_s(X,Y) = H_s(X) + H_s(Y) - H_s(X,Y)$$

Combining Eqs(5) and (6), the computational function for SEMI is shown in Eq. (7):

(7) $$I_s(X,Y) = -\sum_{i=1}^{t} \frac{d_i^{int}}{d_i^{ext}} p_i \log_2(p_i) - \sum_{j=1}^{o} \frac{d_j^{int}}{d_j^{ext}} p_j \log_2(p_j) + \sum_{i=1}^{t} \sum_{j=1}^{0} \frac{d_{ij}^{int}}{d_{ij}^{ext}} p_{ij} \log_2(p_{ij})$$

The spatial entropy-based MI is a special case of the original MI with spatial configuration. Eq. (6) keeps the same form of Eq. (2), so that SEMI also measures the dependence between two images. However, instead of using only one dimensional histogram (counting the number of pixels belonging to each slot), SEMI has the ability to analyze the dependence between the band image and the reference image considering the pixels' spatial distribution. The higher the SEMI value, the more relevant the band image is to the reference image. Therefore, SEMI is used as the new band selection measurement.

2.3. Band Selection Algorithm. Selecting band images with high SEMI values is the general band selection criterion in this paper. However, band selection only based on this criterion will result in high correlation inside the selected band images. Since the reflectance of the earth surface to a certain spectral band is very similar to that of a close spectral band, high correlation exists among neighbor band images. The observation is that only with the SEMI criterion, a series of consecutive band images with high correlation are always selected. To avoid that, another parameter, called band distance, is defined:

Definition 5: Band distance is defined as the absolute difference between two band image indexes:

$$(8) \qquad\qquad \eta = |bandindex_i - bandindex_j|$$

In order to avoid image correlation, each pair of band images selected should have a minimum η threshold. The pseudo code of a new band selection algorithm is designed considering the high SEMI value and low correlation, as shown in Algorithm. 1.

Algorithm 1 BandSelection(CL, mi, num, η)

Require: :

 (1)CL: the candidate band image list
 (2)mi: the vector of SEMI values of all the candidate band images
 (3)num: the number of selected band images
 (4)η: the minimum band distance threshold

1: **Initialize** candidate list $CL = \{1, 2, \cdots, n\}, S = null$
2: $maxBand$=MaxSEMIIndex(CL,mi) and add $maxBand$ into S
3: Remove $maxBand$ from CL
4: **while** S.length$<num$ **do**
5: $maxBand$=MaxSEMIIndex(CL,mi)
6: $flage$=true
7: **for** each band in S **do**
8:   **if** $(|S - maxBand| < \eta)$ **then**
9:    $flag = false$
10: **if** $(flag)$ **then**
11:   Add $maxBand$ into S
12:   Remove $maxBand$ from CL
13: **else**
14:   Remove $maxBand$ from CL
15: **return** S

The band selection algorithm, as in Algorithm. 1, takes four arguments: CL is the candidate band image list, mi is the vector containing the SEMI values of all the candidate band images, num is the number of selected band images and η is the minimum band distance threshold. MaxSEMIIndex(CL, mi) is a function which returns the band image index with the highest SEMI value in CL.

As in Algorithm. 1, the algorithm starts by selecting the band image p with the highest SEMI value from the candidate band image list CL. Then p is removed from CL and added to the selection list S. In the next iteration, the algorithm selects another band image q with the highest SEMI value from CL and removes it

as well. If the band distance between q and each band image in the selection list, e.g., only p in the first iteration, is not smaller than η, q is added to the selection list; otherwise, it will be discarded. This process is repeated until the number of band images in the selection list equals to num. In the end, num band images from the candidate list with the highest SEMI value and low correlation level will be selected.

2.4. SVM Supervised Classification. Support vector machine (SVM) is a supervised classification method [11]. Recent works have shown its effectiveness in land use classification for remote sensing images [8, 13]. In this paper, we will apply SVM for the land use classification to demonstrate the proposed band selection method.

The standard SVM is a binary linear classification method that classifies the input data into two groups by constructing a hyperplane; i.e., two classes locate on the two sides of the hyperplane. Intuitively, a good classification result is achieved by constructing the hyperplane that has the maximum distance to the data located on the margin of the two classes.

For linear inseparable problems, a kernel function is built to map the input space into a higher dimensional space where the linear SVM separation is possible. Different kernel functions exist, including polynomials, radial basic functions (RBF) and hyperbolic tangent [11]. In this paper, we select the most widely used RBF as the kernel function.

One difficulty in remote sensing image classification is the number of the target classes, i.e., the land types. Since the standard SVM is a binary separator, a multiclass SVM scheme, named one-versus-all, is used. The one-versus-all scheme involves the division of an N classes dataset into N two-class cases. For example, given three target land types in remote sensing images, including wood, grass and water area, the one-versus-all scheme will build three SVMs. Each SVM is responsible for one class in which the classification is effected. In other words, three SVM will be responsible for classifying wood against non-wood, grass against non-grass, and water against non-water area, respectively.

In order to test the proposed band selection method, we will use the band selection results as the training inputs to SVM for the land use classification, and the one-versus-all SVM scheme with the RBF kernel function is applied.

3. Experiment

The proposed band selection method is demonstrated using the AVIRIS 92AV3C dataset [2]. All methods are implemented in the Matlab R2009a and experiments are performed on a Dell Optiplex 960 desktop (3G CPU & 4G RAM) using the Windows XP operating system.

3.1. Data Description. The AVIRIS 92AV3C dataset [2] is a public hyperspectral dataset, which is acquired over a test site called Indian Pine in northwestern Indiana [7]. The AVIRIS sensor collects nominally 224 bands of data. Among them, four contain only zeros and are discarded. Therefore, 220 bands from the 92AV3C dataset are used for the experiments.

Each of the 220 band images is of the size of 145×145 pixels. Suggested by the National Aeronautics and Space Administration (NASA), RGB visualization of 92AV3C data uses bands (50, 20, 10) [1], as shown in Figs.2 (a) (b) and (c), respectively, and the final visualization is shown in Fig.2 (d).

The 92AV3C dataset is accompanied with a reference map, indicating the ground truth, as shown in Fig.3. Around 49% pixels are grouped into 16 different classes.

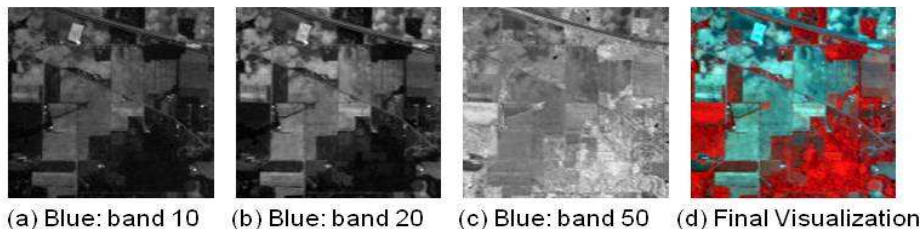(a) Blue: band 10   (b) Blue: band 20   (c) Blue: band 50   (d) Final Visualization

FIGURE 2. 92AV3C data visualization.

As to the remaining pixels, it is difficult to group them into any of the existing class, and they are identified as the background. In the following experiment, 16 labeled classes (excluding the background) are used as ground truth to evaluate the supervised classification accuracy.
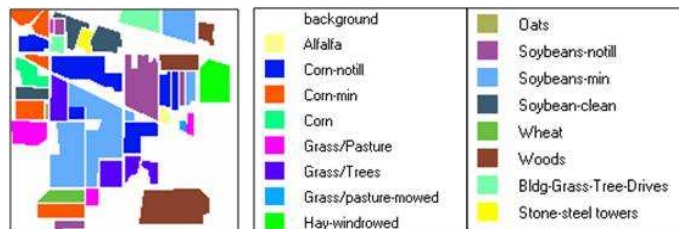

FIGURE 3. Ground truth for 92AV3C dataset.

3.2. Spatial Entropy-based Mutual Information. This experiment is to evaluate the effectiveness of using the spatial entropy-based mutual information (SEMI) as the band selection criterion. SEMI is computed between each of 220 band images in the 92AV3C dataset and the ground truth image, and the result is shown in Fig.4.
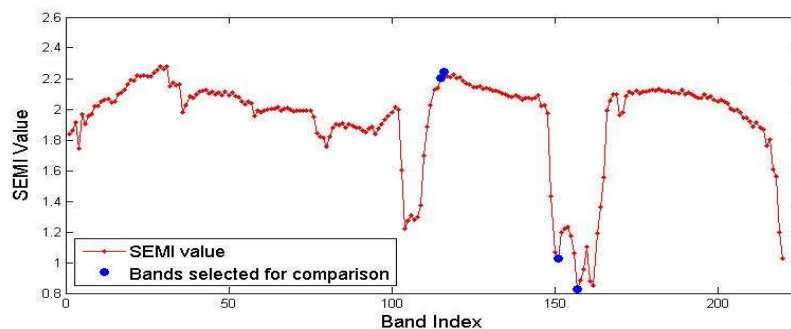

FIGURE 4. SEMI between 220 band images and the ground truth.

SEMI measures the relevant information from the band image to the reference image considering both spatial and non-spatial attributes. The higher the SEMI value they share, the more relevance the band image is to the reference image. For comparison, two band images, (115 and 116), with high SEMI value and two images, (151, 157), with a low SEMI value, are shown in Fig.5. Observation is
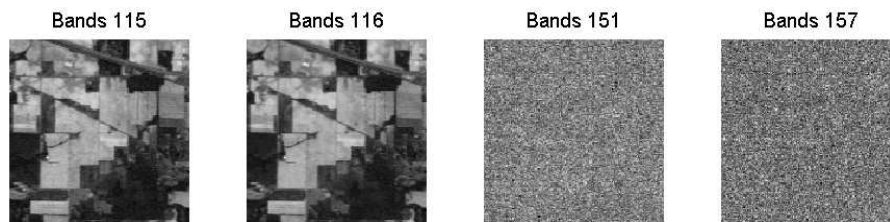
FIGURE 5. Comparison between images with high and low SEMI values: bands 115 and 116 are with SEMI value while bands 151 and 157 are with low SEMI value

that a distinct land pattern can be identified from bands 115 and 116 with a high SEMI value shared with the reference image. On the other hand, bands 151 and 157, with a low SEMI value, are contaminated by the atmosphere water absorption expressing no valuable information for the land use classification. Therefore, it is reasonable to select band images with higher SEMI for the purpose of land use classification.

3.3. Band Selection Result. The proposed band selection algorithm takes four arguments: the candidate list including 220 band images (CL), vector of SEMI values (mi), and number of selected bands (num), minimum band distance threshold (η). From the previous step, the SEMI of 220 band images has been computed. As a case study, we set num to 20, and η to 7. Using the band selection algorithm shown in Algorithm. 1, the 20 selected band images are visualized in Fig.6. 20 bands are almost evenly distributed in the high SEMI value range and of a reasonable band distance from each other. Additionally, it is evident that 20 selected images contain a distinct land pattern and no one is severely contaminated with atmosphere water absorption.
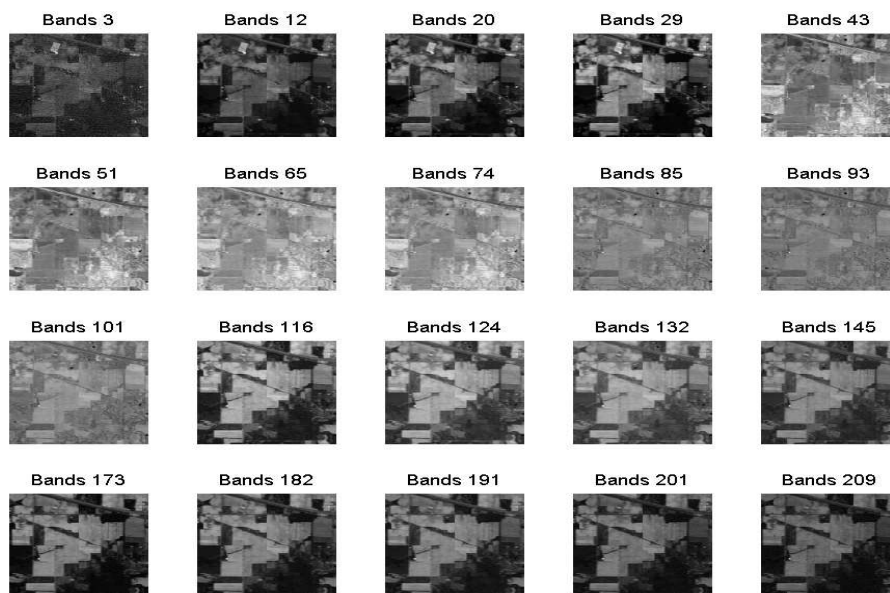

FIGURE 6. Visualization of the 20 selected band images

3.4. Supervised Land Use Classification. To evaluate the band selection results, 20 selected images are used for the supervised land use classification. Several preprocessing steps are performed for this experiment: First, the reference image, from Fig.3, contains 16 predefined classes and 1 background class. Since the areas locating in the background class have ambiguous classification and they do not belong to any of the 16 classes, we remove the pixels belonging to it, which leaves 10,366 pixels for the remaining 16 classes. Second, SVM is a supervised classification method which requires training data to reach a good classifier. The pixels from each class are randomly separated into 60% and 40% as the training and testing data, respectively. For our case study, 6,220 pixels form the training data and the other 4,146 pixels form the testing data.

To classify 16 classes, 16 SVMs are built with applying the one-versus-all scheme and each of them is responsible for one class classification. In the beginning, 6,220 pixels paired with the corresponding pixels from the ground truth are used to train the SVM. Then, 4146 testing pixels are used to evaluate the SVM of the classification accuracy. The result is that 5,706 out of 6,220 pixels from the training data are of correct classification reaching the classification accuracy of 91.7%. For the testing data, the classification accuracy is 88.9%, and the general classification accuracy for both training and testing data is 90.6%.

TABLE 1. Parametric evaluation

Parameters	num	2	5	10	15	20	25	30
	η	50	25	10	9	7	5	4
Classification Accuracy		51.3%	75.3%	84.9%	88.7%	90.6%	93.5%	95.3%

Additionally, a parametric test is performed for the proposed band selection method. With different selection parameters, seven sets of band images are selected for the SVM land use classification. Table.1 shows the parameter configuration and the corresponding classification accuracy, which demonstrates that the classification accuracy increases monotonously with the number of band images selected. For example, the classification accuracy is only 51.3% with two selected band images and it reaches to 95.3% with 30 band images selection. Also, Fig.7 visualizes the classification results and compares them with the ground truth. Visual observations can also tell the constant classification accuracy improvement with the number of band image increases.

3.5. Comparison with Shannon Entropy Based Band Selection Method. The third experiment is to compare the proposed spatial entropy-based band selection method with the previous Shannon entropy-based one. In [7], the Shannon entropy-based MI is applied as the band selection criterion and band images with higher MI values will be chosen.

In order to make a fair comparison, we use the same configurations for the band selection algorithm, as in Algorithm. 1, and the supervised classification. Specifically, two band selection methods take the same arguments: 220 band images from the 92AV3C dataset as CL, the same number of band image selection num and the same minimum band distance threshold η. The only difference is the way to measure the mutual information mi. In comparison of the band selection methods, mi is referred to the Shannon entropy-based mutual information, while in the proposed band selection method, mi is from the spatial entropy-based mutual information (SEMI). Besides, the selected band images are separated with the same portion of training and testing data for the SVM land use classification.

(2 Bands, 51.3%) (5 Bands, 75.3%) (10 Bands, 84.9%) (15 Bands, 88.7%)

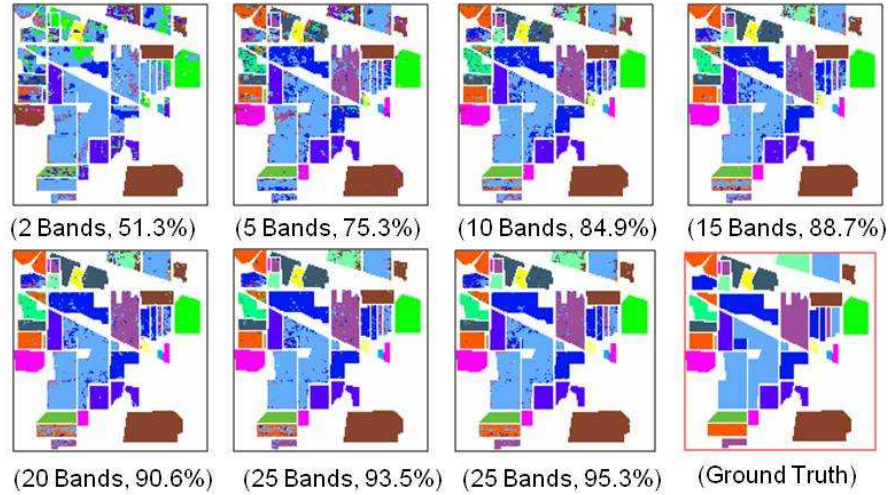(20 Bands, 90.6%) (25 Bands, 93.5%) (25 Bands, 95.3%) (Ground Truth)

FIGURE 7. Supervised classification results comparison by using different number of selected bands
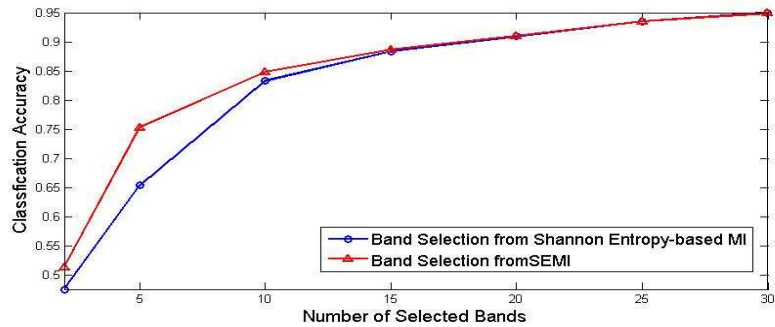

FIGURE 8. Classification accuracy comparison between proposed method and previous Shannon entropy-based method

Fig.8 compares the classification accuracy. The triangle marked curve is from the spatial entropy-based band selection method while the circle marked one is from the Shannon entropy-based method. Visual observation shows that the classification accuracy using the spatial entropy-based method is generally higher than the Shannon entropy-based one. Especially, with a smaller number of band images selection, the spatial entropy-based band selection method has apparent advantage. For example, with 5 band image selection, the classification accuracy using the spatial entropy-based band selection results is 75.3% which surpasses the Shannon entropy-based method, 65.1%, by more than 10 percent.

4. Conclusion and Future Work

This paper proposes a new spatial entropy-based method for the hyperspectal band selection. Integrating the spatial entropy with the traditional mutual information measurement, a new spatial entropy-based mutual information (SEMI) function is derived and introduced as the band selection measurement. Specifically, the higher SEMI value the band image shares with the ground truth, the more relevant it is related to the classification. Furthermore, this paper develops a new band

selection algorithm taking into consideration of both SEMI values and band image correlation. Experiments demonstrate that the proposed band selection method offers higher classification accuracy for land use datasets.

The new spatial entropy-based band selection method still needs improvement: First, the current version of the SEMI function is computationally expensive. In the future, we need to modify the methods and design the data structure to improve the efficiency. Second, the proposed method will be applied to other large remote sensing datasets.

References

[1] AVIRIS Starting Guide, http://aviris.jpl.nasa.gov/html/AVIRIS_for_dummies.html, 2010
[2] 92AV3C Data Source, ftp://ftp.ecn.purdue.deu/biehl/MultiSpec, 2010
[3] A. Martinez-Uso, F. Pla, P. Garcia-Sevilla, and J.M. Sotoca, Automatic band selection in multispectral images using mutual information-based clustering, Proc. 11th Iberoamerican Congr. Pattern Recog., Cancun, Mexico, (2006) 644-654.
[4] C. Claramunt, A spatial form of diversity, Proc. of COSIT2005, Ellicottville, USA, (2005) 218-231.
[5] C. Conese and F. Maselli, Selection of optimum bands from TM scenes through mutual information analysis, ISPRS J. Photogramm, 48(3) (1993) 2-11.
[6] D. Landgrebe, Hyperspectral image data analysis as a high dimensional signal processing problem, IEEE Signal Proc. Mag., 19 (2002) 17-28.
[7] B.F. Guo, R.I. Damper, S.R. Gunn, and J.D.B. Nelson, A fast separability-based feature-selection method for high-dimensional remotely sensed image classification, Pattern Recognit., 41(5) (2008) 1653-1662.
[8] C. Huang, L.S. Davis and J.R.G. Townshend, An assessment of support vector machines for land cover classification, Int. J. Remote Sens., 23(4) (2002) 725-749.
[9] AVIRIS General Overview, http://aviris.jpl.nasa.gov/html/aviris.overview.html, 2010
[10] P. Bajcsy and P. Groves, Methodology for hyperspectral band selection, Photogram. Eng. Remote Sens., 70(7) (2004) 793-802.
[11] S. Haykin, Neural networks: A Comprehensive Foundation, 2nd Edition, Pearson Education, 2004.
[12] W. R. Tobler, Computer movie simulating urban growth in Detroit Region, Econ. Geogr., 46(2) (1970) 234-240.
[13] D. Tuia, F. Pacifici, M. Kanevski, and W. J. Emery, Classification of very high spatial resolution imagery using mathematical morphology and support vector machines, IEEE Trans. Geosci. Remote Sens., 47(11) (2009) 3866-3879.
[14] X. Li and C. Claramunt, A Spatial Entropy-Based Decision Tree for Classification of Geographical Information, Trans. in GIS, 10(3) (2006) 451-467

Department of Geomatics Engineering, Schulich School of Engineering, University of Calgary, 2500 University Drive NW, Calgary, Alberta T2N 1N4
  E-mail: {baiwang,xcwang}@ucalgary.ca
  URL: http://www.ucalgary.ca/wangx/

Center for Computational Geoscienes and Mathematics, Faculty of Science, Xi'an Jiaotong University, Xi'an,710049, P.R. China and Department of Chemical and Petroleum Engineering, Schulich School of Engineering, University of Calgary, 2500 University Drive NW, Calgary, Alberta T2N 1N4
  E-mail: zhachen@ucalgary.ca
  URL: http://www.ucalgary.ca/reservoir_simulation/