ISSN Online: 2821-1693

Transactions on Machine Intelligence

Journal Homepage: https://tmachineintelligence.ir/index.php/TMI

Two New Indices for Unsupervised Dimensionality Reduction of Hyperspectral Data

B. Asghari Beirami[1,*], M. Mokhtarzade[2]

[1] Ms.c of Remote Sensing, Department of Photogrammetry and Remote Sensing, K.N Toosi University of Technology, Tehran, Iran
[2] Associate Professor, Department of Photogrammetry and Remote Sensing, K.N Toosi University of Technology, Tehran, Iran

ARTICLE INFO

Article History:
Received 29 June 2022
Received in revised form
4 October 2022
Accepted 11 November 2022
Available online 12 November 2022.

Keywords:
Dimensionality reduction, PCA, Wavelet, Spectral signature curve (SSC), Trapezoidal numerical integration, L2-norm

ABSTRACT

Hyperspectral imaging, with hundreds of spectral bands provide a rich data source for thematic mapping. However, the high dimensionality of these images is a challenge which is commonly treated by dimensionality reduction techniques. In this paper, a pixel-based hyperspectral dimensionality reduction technique is proposed in which the spectral signature curve (SSC) is divided into segments with equal lengths. Afterwards two new mathematical indices, which are based on integral and L2-norm, are proposed to transform each SSC segment into a new feature of the reduced feature vector. The proposed method considers the ordinance of SSC and, unlike PCA, can be applied individually to all pixels independently and simultaneously which means the method is applicable in parallel processing. The proposed data reduction technique was applied to two well-known agricultural hyperspectral scenes and was compared to wavelet and PCA data reduction techniques. The obtained results proved the efficiency of the proposed method from both classification accuracy and processing time aspects.

1. INTRODUCTION

Hyperspectral sensors are modern tools for earth phenomena observation which acquire data in very narrow partitions of electromagnetic spectrum and form a data cube. Spectral Signature Curve (SSC) is the plot of the sequence of reflectance values against wavelengths (or band numbers). Hyperspectral sensors approximate SSCs appropriately which leads to better identification of ground materials.

Curse of dimensionality and huge processing time are serious challenges in processing of hyperspectral data, and thus, data reduction techniques are commonly applied to these images. Generally, feature reduction methods are divided into two groups, namely feature selection and feature extraction. Feature selection methods try to find

[*] Corresponding author: Behnam.asghari1370@gmail.com
Ms.c of Remote Sensing, Department of Photogrammetry and Remote Sensing, K.N Toosi University of Technology, Tehran, Iran

http://dx.doi.org/10.47176/TMI.2022.21

optimal features among original set of features while feature extraction methods attempt to find a transformed space in which optimal features are simply distinguished.

Among the most common feature extraction methods one can point to principal component analysis (PCA) [1], minimum noise fraction (MNF) [2] as unsupervised techniques while DAFE [3], NWFE [4], DBFE [5] are the most famous supervised ones.

The aforementioned feature extraction methods are performed on the whole image at once and do not consider SSC ordinance which is a rich source of information. On the other hand, there are some pixel based techniques in the literature that consider the geometric shape and ordinance of SSC. These techniques, which are the group that our method belongs to, are reviewed below.

Fractal dimension of SSCs, as a new way of hyperspectral feature extraction was introduced in [6]. This technique was later developed [7] in which fractal dimensions were derived for each segments of SSCs. Also, [8] studied the effect of adjacent segments overlap on quality of reduced features in fractal dimension based dimensionality reduction. Rational function curve fitting concept was proposed in [9] as a new feature extraction method. In their method polynomial coefficients in the numerator and denominator of the fitted curve are considered as new extracted features. In [10] an automatic method is proposed for dimension reduction of hyperspectral data which is based on wavelet decomposition of SSCs. Their proposed method outperforms PCA both from accuracy and processing time aspects. The general description of their algorithm is as follow:

1) The Daubechies-4 wavelet is used to decompose the 1-D signal corresponding to each pixel in a hyperspectral image. The Mallat algorithm applies two low-pass and high-pass filters to the signal, followed by dyadic decimation, which removes every other part of the signal and thereby cuts its overall length in half. The aforementioned approach is repeated recursively by applying the same procedure to the results of the filter sub-bands to produce an ever smoother version of the original vector.

2) The inverse discrete wavelet transform is used to recover an estimate of the original spectral for each hyperspectral pixel (IDWT). The level of decomposition required for a given pixel is the one that produces an acceptable correlation with the original signature.

3) The number of decomposition levels (L) is automatically determined by combining the findings from all pixels and is considered the lowest needed level after rejecting the outliers.

4) The reduced output data are created by decomposing all pixels to level L using the number of L.

The proposed method of this paper also considers the geometric shape and ordinance of SSC. The main contribution of the current paper is the presentation of two new indices to effectively reduce the dimensionality of hyperspectral images with appropriate processing efforts. This method divides the spectral signature curve (SSC) into equally length segments and then computes two new mathematical indices, namely trapezoidal numerical integration and normalized square of L2-norm, in each segment. Another advantage of the proposed method is that it can be applied pixel by pixel and does not need to transform the whole data to a new space simultaneously as it is the case in methods like PCA. It also considers the spectral ordinance of SSC and has capability of parallel processing.

The proposed data reduction technique was applied to two well-known medium spatial resolution Indian pines and high spatial resolution Salinas-A agricultural hyperspectral scenes and results were compared to wavelet and PCA data reduction techniques through overall classification accuracy of KNN and SVM classifiers.

In the second part we explain the methodology of the proposed dimensionality reduction method which is followed by the introduction of the study data sets. Afterwards the implementation results are presented and evaluated. The results are also compared to wavelet and PCA methods in terms of classification accuracies and processing times. Conclusions are presented in the last section.

2. METHODOLOGY

In this paper a pixel-based method, which is capable to be implemented in parallel processing schemes, is proposed to reduce the dimensionality of hyperspectral data. Figure 1 shows the structure of the proposed method.

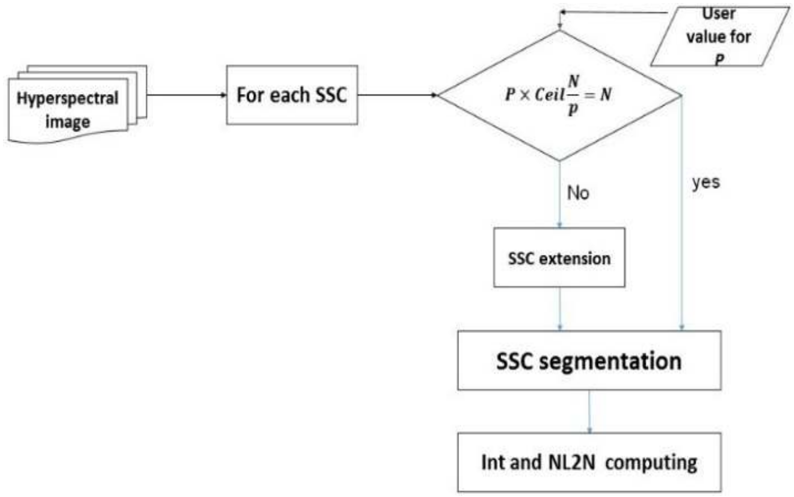

Fig. 1. Workflow of proposed dimensionality reduction method

According to the above figure, the proposed method consists of two main stages which are 1- SSC segmentation and 2-feature extraction via transformation of each segment into a new feature based on some mathematical indices. These stages are described in the following.

2.1. SSC segmentation

Given N, which is the number of hyperspectral image bands, user determines the desired number of reduced bands (P). It is expected to have P segments with the length of $\text{Ceiling}(N/P)$. However, when $\text{Ceiling}(N/P) \times P \neq N$ the last segment cannot have the full length of $\text{Ceiling}(N/P)$ and SSC should be extended.

Many different methods of extension can be applied such as zero extension, replicate extension, circular extension and symmetric extension. Based on our experiments, best results were achieved via symmetric extension mode.

2.2. Feature extraction

This section aims to assign an index to each SSC segment which is regarded as a component of the reduced feature vector. To do so, two indices are introduced in this paper, namely *Int* index and *NL2N* index.

Int index is simply the area under the sub-curve of SSC segment which is determined via numerical techniques. Suppose that there are $T+1$ evenly spaced points in each segment represented as $a = x_1 < x_2 < ... < x_{T+1} = b$. The first index can be computed via equation 1.

$$\int_a^b f(x)dx = \frac{b-a}{2T}\sum_{n=1}^{T}(f(x_n) + f(x_{n+1})) = \frac{b-a}{2T}[f(x_1) + 2f(x_2) + \cdots + 2f(x_T) + f(x_{T+1})] \quad (1)$$

In the above equation, x_i and $f(x_i)$ values represent band numbers and their reflectance value for all the bands of that segment.

To compute the second index, called *NL2N*, lets define a vector in each segment as (2)

$$y = [f(x_1)\ f(x_2)\ \dots\ f(x_T)\ f(x_{T+1})]^{\text{Transpose}} \quad (2)$$

In this case, *NL2N* index can be computed as the L2-norm of y (equation 3)

$$\frac{||y||^2}{T+1} \quad (3)$$

3. EXPERIMENTAL RESULTS

3.1. Data set

The Indian Pines scene was captured by the AVIRIS sensor in 220 spectral bands from a mixed forest/agricultural 145 145 pixels from the Indian Pine Site. This data collection has a spatial resolution of 20 m and band wavelengths ranging from 400 to 2500 nm with a resolution of 10 nm. This investigation employed 200 bands after deleting water absorption bands. There are 16 different land covers in this landscape. Because of the comparable reflectance amongst key classes, classifying this scene has proved difficult.

The 224-band AVIRIS sensor captured this high-resolution Salinas-A picture over Salinas Valley, California (3.7-meter). The image has a resolution of 8683 pixels. We deleted the 20 water absorption bands, in this case bands [108-112], [154-167], and 224, as we did in the Indian Pines scene. This image, which contains 6 vegetal classes, was only available as at-sensor radiance data.

3.2. Results and discussion

The accuracy of a classification map created by the reduced data is commonly used to evaluate data reduction techniques. Two well-known classification algorithms, K-Nearest Neighbor (KNN) and support vector machines (SVM), were used in this study. A pixel is classified in the KNN classifier technique by a majority vote on the labels of neighboring pixels in feature space, with the pixels given to the class most prevalent among its k nearest neighbors. If k = 1, for example, the pixel is simply allocated to the class of its single nearest neighbor. SVM is a supervised learning method that determines the best hyperplane to optimize the distance between the closest training sample and the separating hyperplane.

SVMs can conduct non-linear classification using kernel functions, implicitly mapping their inputs into high-dimensional feature spaces, in addition to doing linear classification. This method aids the algorithm in locating the maximum-margin hyperplane in a converted space.

For the sake of comparison, PCA and wavelet as two other well-known data reduction techniques were also implemented and were assessed in the same manner. The performance of wavelet based date reduction highly depends on the kind of mother wavelet. According to [10] in this study Daubechies-4 was used. In this technique, 2 to 5 decomposition levels of wavelet were implemented leading to different number of reduced feature vector. Number features in tables 1 and 2 are set based on decomposition levels of wavelet method.

KNN classifier requires us to set the number of K-neighbors in advance. In this paper, different K values from one to ten were examined to locate its optimum value. Euclidean distance, as the most common case in the literature, was also set as the distance measure.

SVM classifier has also some adjusting parameters which should be set in advance. The most important one is, perhaps, the kernel function. Different kernel functions such as linear, Gaussian, quadratic and cubic was examined and because of better performance of cubic kernel, in this paper, this kernel function was applied.

The other parameters of SVM, namely kernel scale (γ) and penalty term (c) were set via cross-validation; a one against-one multi-class classification strategy was adopted, (γ) parameter was determined automatically at optimal and penalty parameter c was tested between $[2^{-4} - 2^{4}]$. Training samples were randomly selected from less than 10% for Indian pine and 1.9 % for Salinas–A and the entire ground-truth maps, except those were which were used for training, were applied as test samples. The results of classification for both data sets are presented in table 1 and 2.

SVM and KNN Classification overall accuracies for Indian pine and Salinas-A with original 200 and 204 spectral bands are 79.14, 68.23 and 98.38, 97.45 respectively. According to tables 1 and 2, classification accuracy decreases for PC features as the number of features increase; this is maybe because of lower level noisy features. Also classification accuracies with wavelet and proposed method are commonly higher than original classification accuracy results for both data sets.

Table 1. Classification accuracy for Indian pine

Classifier	Reduction method	Number of features (based on decomposition levels of wavelet)			
		55	31	19	13
SVM	Wavelet	79.60	78.37	77.66	75.21
	PCA	70.42	71.26	73.87	74.20
	Proposed method (NL2N)	79.83	79.15	79.28	79.21
	Proposed method (Int)	80	79.41	80.01	78.91
KNN	Wavelet	69.59	68.99	65.86	61.41
	PCA	59.84	61.88	63.88	63.88
	Proposed method (NL2N)	72.21	72.67	72.28	71.33
	Proposed method (Int)	71.58	71.78	71.95	70.72

Table 2. Classification accuracy for Salinas-A

Classifier	Reduction method	Number of features (based on decomposition levels of wavelet)			
		56	31	19	13
SVM	Wavelet	98.68	98.55	97.66	98.36
	PCA	84.24	90.76	94.21	95.84
	Proposed method (NL2N)	98.72	98.74	98.67	98.60
	Proposed method (Int)	98.68	98.72	98.70	98.55
KNN	Wavelet	98.02	98.02	97.54	95.90
	PCA	44	54.76	69.32	77.53
	Proposed method (NL2N)	98.65	98.68	98.63	98.48
	Proposed method (Int)	98.59	98.57	98.38	98.28

The obtained results also prove the robustness of the proposed method versus the number of features as the overall accuracy does not vary considerably for different number of features. Better performance of proposed method can be assigned to its capability in preserving the complexity, ordinance, low pass, high pass and discriminative parts of SSCs. To examine this claim, figure 2 is prepared in which the reduced feature vector of the proposed method and its opponents (i.e. PCA and wavelet) are shown versus the original spectral curve. In wavelet method, an exhaustive effort is needed to select the best mother wavelet. This issue, however, is relaxed in our proposed method. As another advantage of the proposed method, one can point to the simplicity of the applied indices which yields to better computational efficiency and makes them appealing to be applied in other techniques.

3.3. Time efficiency of the proposed method

In this subsection time efficiency of the proposed method is investigated and the results are compared to wavelet as better competitor method. Both methods were implemented in MATLAB 2016b software at personal computer with Intel® Core™ 2 Quad CPU 2.66 GHz with 4.00 GB RAM and 64 bit operating windows 10. Wavelet toolbox of Matlab was adopted for generating wavelet features. Spent seconds for dimensionality reduction are illustrated in figure 3.

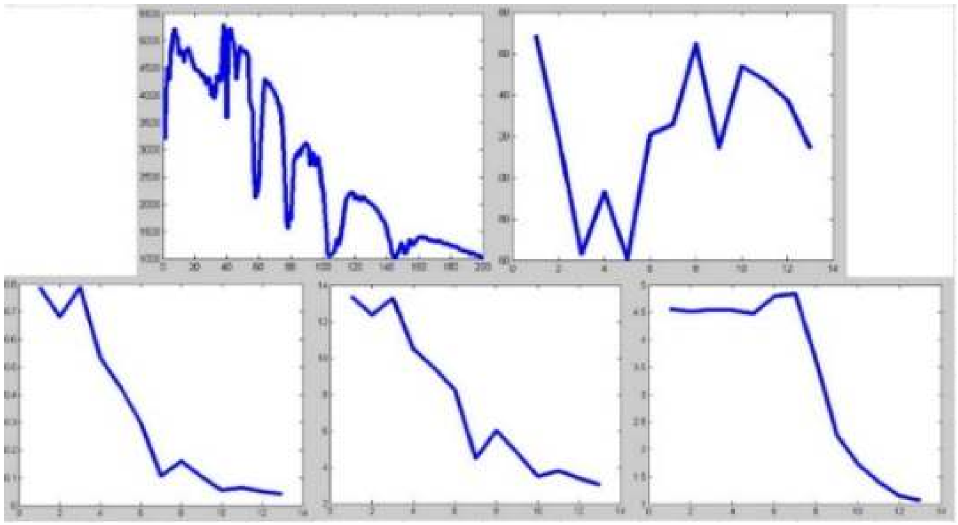

Fig. 2. Original and dimensionally reduced SSC for pixel (1, 1) Above. Left) original SSC for a sample pixel at Indian pine data set- above. Right) dimensionally reduced data with 13 first PCA- below. Dimensionally reduced data based) left. N2LN-midle.INT index-right. Wavelet

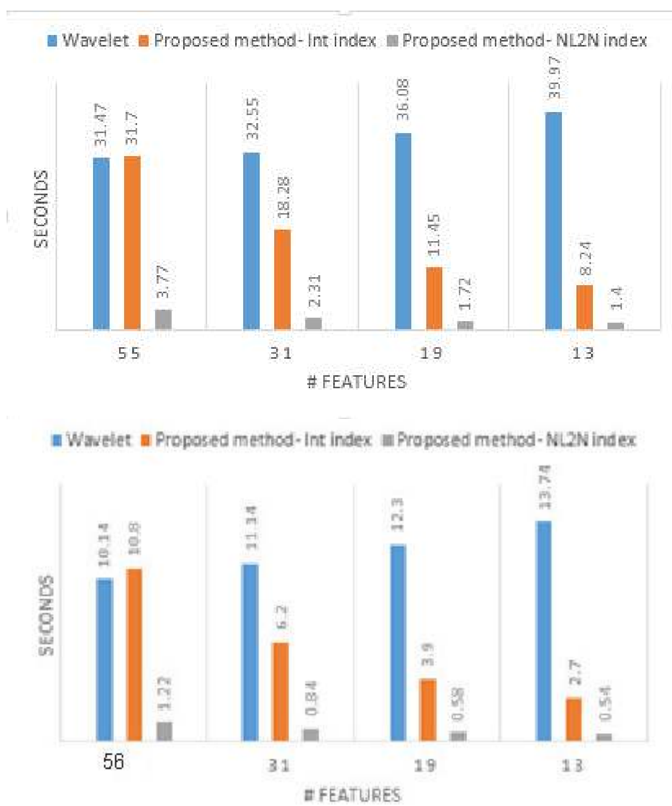

Fig. 3. Spent seconds for dimensionality reduction – above) Indian pine below) Salinas-A

According to figure 3, wavelet processing time increases proportional to the number of its decomposed levels. This is because the algorithm needs additional works [10]. On the other hand, in the proposed method the spent processing time decreases with increasing compression level. Among the two proposed indices, *NL2N* index is much more efficient from the processing time point of view.

4. CONCLUSION

In this paper we proposed a new simple dimensionality reduction method which is mainly targeted for hyperspectral images. In that method, spectral signature curve (SSC) is divided into equally length segments and the information content of each segment is reduced to a representative index. Two indices, namely trapezoidal numerical integration (*INT*) and normalized square of L2-norm (*NL2N*), were introduced and the results were compared from accuracy and processing time aspects. The results were also compared to PCA and wavelet date reduction techniques. Turning to the comparison, the proposed method yielded to a reduced feature which could provide the best classification accuracy. In that regards, both indices led to comparable results which were better than wavelet for at average about 2 % for first data set and about 0.7 % for second data set and outperformed PCA for at average about 7 % for first data set and 7.5 % for second data set. From the computational point of view, the proposed method revealed very appealing as it was faster that wavelet, especially in cases that higher decomposition levels are set in the wavelet analysis. The proposed method exploits the information content existing in ordinance of SSCs. It also benefits from being applicable in parallel processing due to its pixel-by-pixel computational nature. The proposed method requires us to set the segmentation number. Automatic selection of this parameter and flexible adaption of segments lengths are the subjects for our more study.

REFERENCES

[1] Lim, S., Sohn, K. H., & Lee, C. (2002). Principal component analysis for compression of hyperspectral images. IGARSS 2001. Scanning the Present and Resolving the Future. Proceedings. IEEE 2001 International Geoscience and Remote Sensing Symposium (Cat. No.01CH37217). IEEE 2001 International Geoscience and Remote Sensing Symposium, Sydney, NSW, Australia. doi:10.1109/igarss.2001.976068

[2] Yang, C., Everitt, J. H., & Johnson, H. B. (2009). Applying image transformation and classification techniques to airborne hyperspectral imagery for mapping Ashe juniper infestations. International Journal of Remote Sensing, 30(11), 2741–2758. doi:10.1080/01431160802555812

[3] Landgrebe, D. A. (2008). Signal theory methods in multispectral remote sensing. Newy York: Wiley-Interscience.

[4] Kuo, B. C., & Landgrebe, D. A. (2004). Nonparametric weighted feature extraction for classification. IEEE Transactions on Geoscience and Remote Sensing, 42(5), 1096-1105.

[5] Lee, C., & Landgrebe, D. A. (1993). Feature extraction based on decision boundaries. IEEE Transactions on Pattern Analysis and Machine Intelligence, 15(4), 388-400.

[6] Ghosh, J. K., & Somvanshi, A. (2008). Fractal-based dimensionality reduction of hyperspectral images. Journal of the Indian Society of Remote Sensing, 36(3), 235–241. doi:10.1007/s12524-008-0024-0

[7] Hosseini, A., & Ghassemian, H. (2012, Μάιος). Classification of hyperspectral and multispectral images by using fractal dimension of spectral response curve. 20th Iranian Conference on Electrical Engineering (ICEE2012, Tehran, Iran. doi:10.1109/iraniancee.2012.6292587

[8] Hosseini, S. A., & Ghassemian, H. (2013, Μάιος). A new hyperspectral image classification approach using fractal dimension of spectral response curve. 2013 21st Iranian Conference on Electrical Engineering (ICEE). Mashhad, Iran. doi:10.1109/iraniancee.2013.6599552

[9] Hosseini, S. A., & Ghassemian, H. (2016). Rational function approximation for feature reduction in hyperspectral data. Remote sensing letters, 7(2), 101–110. doi:10.1080/2150704x.2015.1101180

[10] Kaewpijit, S., Le Moigne, J., & El-Ghazawi, T. (2003). Automatic reduction of hyperspectral imagery using wavelet spectral analysis. IEEE transactions on geoscience and remote sensing: a publication of the IEEE Geoscience and Remote Sensing Society, 41(4), 863–871. doi:10.1109/tgrs.2003.810712