

Adaptive Lifting Transform for Classification of Hyperspectral Signatures

Rajesh Agrawal¹, Narendra Bawane²

¹Department of Electronics, GHRC, Nagpur, India

²RTM University, Nagpur, India

Email: rkdhule@yahoo.co.in, narendrabawane@rediffmail.com

Received 30 April 2015; accepted 21 June 2015; published 24 June 2015

Copyright © 2015 by authors and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Supervised classification of hyperspectral images is a challenging task because of the higher dimensionality of a pixel signature. The conventional classifiers require large training data set; however, practically limited numbers of labeled pixels are available due to complexity and cost involved in sample collection. It is essential to have a method that can reduce such higher dimensional data into lower dimensional feature space without the loss of useful information. For classification purpose, it will be useful if such a method takes into account the nature of the underlying signal when extracting lower dimensional feature vector. The lifting framework provides the required flexibility. This article proposes the adaptive lifting wavelet transform to extract the lower dimensional feature vectors for the classification of hyperspectral signatures. The proposed adaptive update step allows the decomposition filter to adapt to the input signal so as to retain the desired characteristics of the signal. A three-layer feed forward neural network is used as a supervised classifier to classify the extracted features. The effectiveness of the proposed method is tested on two hyperspectral data sets (HYDICE & ROSIS sensors). The performance of the proposed method is compared with first generation discrete wavelet transform (DWT) based feature extraction method and previous studies that use the same data. The indices used for measuring performance are overall classification accuracy and Kappa value. The experimental results show that the proposed adaptive lifting scheme (ALS) has excellent results with a small size training set.

Keywords

Adaptive Lifting Scheme, Pattern Classification, Hyperspectral Image, Feature Extraction, Neural Network

1. Introduction

Hyperspectral data set consists of hundreds of images corresponding to different wavelengths for the same area

on the surface of the Earth [1]. The resulting higher dimensional feature vector (hyperspectral signature) causes the problem of dimensionality for conventional classifiers [2]. That is, the training data set size must increase exponentially with the dimension of the feature vector. However, practically, limited numbers of labeled pixels are available and therefore we cannot reliably estimate the classifier parameters [3]. Thus, it is essential to have a feature reduction method that can map such higher dimensional data into lower dimensional feature space without the loss of useful information.

A feature reduction approach based on feature-extraction projects the original feature space onto the lower dimensional space by transformation. Recently, wavelet-based multiresolution analysis has been the widely used feature-extraction method in signal processing [4]. The wavelets provide an optimal representation for many signals containing singularities [5]. Therefore, it is feasible to detect the singularity of reflectance spectra by using the wavelet transform. Hsu *et al.* have proposed the discrete wavelet transform (DWT) and wavelet packet transform to extract lower dimensional spectral features for classification of hyperspectral images [6]. The DWT is proposed for feature extraction and signature classification; however, the authors mention that DWT is not shift invariant [7]. A small shift in singularity causes large variation in wavelet coefficient oscillation pattern. Around singularities, the wavelet coefficients tend to oscillate and may have a very small or even zero value. Such behavior of wavelet coefficients makes the singularity detection very difficult [8].

Due to its fixed filter bank structure, a wavelet does not always capture all the transient features of the input signal. This may result in lower classification accuracy. The conventional convolution-based implementation of the DWT has high computational and memory cost. For classification purpose, it will be useful to have a multi-resolution tool that takes into account the nature of the underlying signal. The lifting framework proposed by Wim Sweldens provides the required flexibility [9].

This article proposes the adaptive lifting wavelet transform to extract the lower dimensional feature vectors for classification purpose. The lifting framework allows the decomposition filter to adapt the input signal so as to improve or leave intact the desired characteristics of the signal. Most adaptive lifting schemes proposed by the researchers have been used for image compression and denoising application: adaptive predict [10] [11] and adaptive update [12] [13]. For lossless image compression, a three-step nonlinear lifting scheme adapts both the operators and results in fewer large detail coefficients [14]. For denoising higher frequency features such as edges, very efficient point wise adaptive wavelet transform is presented [15]. The transform uses intersection of confidence intervals rule to determine the filter support for each sample independently. For epileptic seizure electrocardiogram (ECG) classification, Subasi and Ercelebbi used lifting-based discrete wavelet transform as a preprocessing method to increase the computational speed. The results are compared with that obtained using first generation wavelets [16].

In the context of remotely sensed images, most of the existing studies propose lifting scheme for lossless coding and for denoising. The multiplicative speckles in synthetic aperture radar (SAR) images were reduced by wavelet transform based on lifting scheme [17]. To enhance the compression efficiency, the design of lifting filter adapted to the signal statistics was proposed by Gouze *et al.* For example, the prediction filter is designed to minimize the variance of the signal [18]. To improve the performance of lossless compression for multichannel image, the blockwise adaption of the coefficients of the predictors of VQLS was proposed [19].

Thus, no literature was found on hyperspectral image classification within lifting framework. In this article, we propose the adaptive update operator that retains the transient features of the hyperspectral signature that helps to improve the classification. The rest of the article is organized as follows. The proposed adaptive update lifting scheme is introduced in Section 2. The experiments and results are presented in Section 3 and Section 4, respectively. Finally, conclusions follow in Section 5.

2. Proposed Lifting Scheme

As the application in hand is to produce the feature vectors from hyperspectral signatures for classification, the objective is to retain the distinctive information of the input signature. The spectral signature of different land cover classes are shown in [Figure 1](#). It is observed that distinctive information lies in signal discontinuities, transitions or edges. Therefore, the proposed lifting scheme must preserve these discontinuities. The proposed adaptive lifting wavelet transform illustrated in [Figure 2](#) consists of three steps described as follows.

2.1. Split (Lazy Wavelet Transform)

It divides the hyperspectral signature (underlying input signal) $X(z)$ into two polyphase components: even

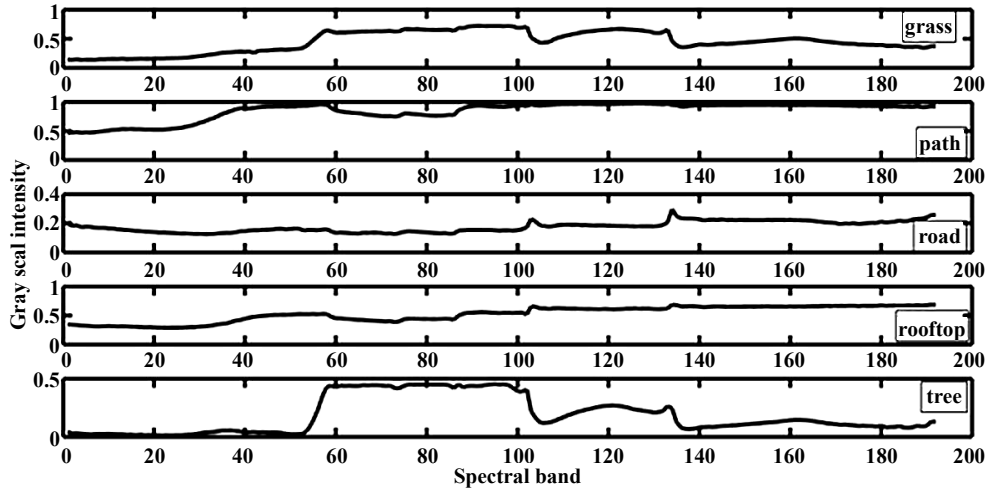


Figure 1. Hyperspectral signature of various landcover classes.

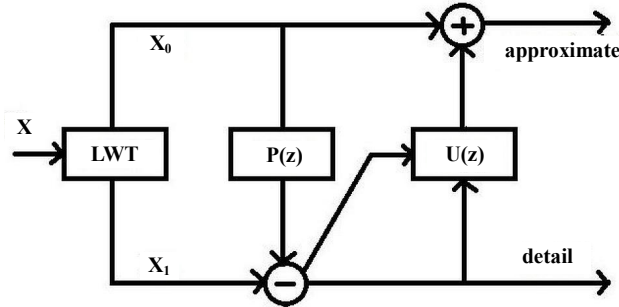


Figure 2. Adaptive lifting scheme structure.

indexed $X_0(z)$ and odd indexed samples $X_1(z)$.

2.2. Fixed Predict

This step produces the detail coefficients $X_1^{new}(z)$. The odd elements $X_1(z)$ are predicted from even elements $X_0(z)$. The difference between an odd sample and prediction gives a detail coefficient. That is, $X_1^{new}(z) = X_1(z) - P(z)X_0(z)$. To build a good predictor one should know the nature of the original signal and the problem in hand. In the case of data compression applications, the goal is to minimize the detail signal. For example, for a triangular signal, detail coefficients can be made zero if prediction is the average of two neighboring even samples. In this algorithm, we use a fixed predict, that is, even sample itself is the best prediction of odd samples. Thus $P(z) = 1$. Therefore we have,

$$X_1^{new}(z) = X_1(z) - X_0(z) \tag{1}$$

In time domain, $x_1^{new}(n) = d(n) = x(2n+1) - x(2n)$. Thus the detail coefficients are computed as the difference between two adjacent samples. The fixed predict $P(z) = 1$ is justified because if the original hyperspectral signature has local coherence, then both even and odd samples would be highly correlated. In that case, the difference $d(n)$ and hence the detail coefficient will be negligible. On the contrary, at discontinuities, even and odd samples are not correlated.

Therefore, the given fixed predictions results in large differences and hence detail coefficients, thereby retaining the distinct information about discontinuities.

2.3. Adaptive Update

Because of subsampling at first stage, even samples $X_0(z)$ are unsuitable for approximate signal. In the third

step, even samples are updated using wavelet coefficients $X_1^{new}(z)$ to produce the approximate signal $X_0^{new}(z)$. That is, $X_0^{new}(z) = X_0(z) + U(z)X_1^{new}(z)$. The update step provides the approximate signal $X_0^{new}(z)$ that will be used as input for the next level of decomposition. The purpose of this update stage is to retain some property (such as mean value) of the original signal over successive decomposition. In our application, the objective is to retain the distinctive information of the input signature. The algorithm must not smooth the discontinuities for the approximate signal. Therefore, depending upon the type of region, the update step varies as follows: At discontinuities when $d(n)$ has large value, the signal is not smoothed but rather kept as it is to retain the information. In the smooth region when $d(n)$ is small, the approximate signal $X_0^{new}(z)$ is computed as the average of odd and even samples. Thus the adaptive update function is as follows.

$$U(z) = \begin{cases} 0.5, & \text{if } |d(n)| < T \\ 0, & \text{if } |d(n)| \geq T \end{cases} \quad (2)$$

Here, T is the threshold computed as: $T = 0.5 * [\max(d(n)) - \min(d(n))]$. The approximate component is expressed as follows.

$$X_0^{new}(z) = \begin{cases} X_0(z) + 0.5 * X_1^{new}(z), & \text{if } |d(n)| < T \\ X_0(z), & \text{if } |d(n)| \geq T \end{cases} \quad (3)$$

In time domain,

$$x_0^{new}(n) = \begin{cases} x_0(n) + 0.5 * x_1^{new}(n), & \text{if } |d(n)| < T \\ x_0(n), & \text{if } |d(n)| \geq T \end{cases} \quad (4)$$

As $d(n) = x_1(n) - x_0(n)$; $x_1(n) = x(2n+1)$ and $x_0(n) = x(2n)$, we have,

$$x_0^{new}(n) = \begin{cases} \frac{x(2n) + x(2n+1)}{2} & \text{if } |d(n)| < T \\ x(2n), & \text{if } |d(n)| \geq T \end{cases} \quad (5)$$

Because of the above adaptive update function, the approximate signal retains the discontinuity information. Thus the proposed lifting scheme takes the average of neighboring samples in the smooth regions of the signal and does not modify the signal at edges to retain the distinctive information useful for classification.

3. Experiments

3.1. Hyperspectral Data Set

To study the effectiveness of the proposed adaptive update lifting scheme for classification, experiments are done on the following hyperspectral data sets. These images are used as it is without any processing.

Washington DC Mall data set is collected by the Hyperspectral Digital Imagery Collection Experiment (HYDICE) system, consisting of 191 spectral channels in the region of the visible and infrared spectrum [20] [21]. For decomposition, the last spectral band is repeated to make the dimension of the signature in terms of power of 2. Throughout the experiments in this study, we used the 256 * 256 portion of the image shown in **Figure 3**. The image has five land cover identified as follows: roof top, grass, road, path, and trees.

University of Pavia data set consists of Pavia University scenes acquired by the Reflective Optics System Imaging Spectrometer (ROSIS) sensor [22]. Out of available 103 spectral bands, only the first 96 bands are used in this study. A 610 * 340 pixels section shown in **Figure 4** is used for experiments. The scene has nine ground cover identified as follows: Asphalt, Meadows, Gravel, Trees, Painted metal sheets, Bare soil, Bitumen, Self-Blocking Bricks, and Shadows.

3.2. Neural Network and Its Topology

In this experiment, a single hidden layer feed forward neural network is used for classification of feature vectors. After training by back propagation algorithm, the network is used as a classifier to classify the whole image. The

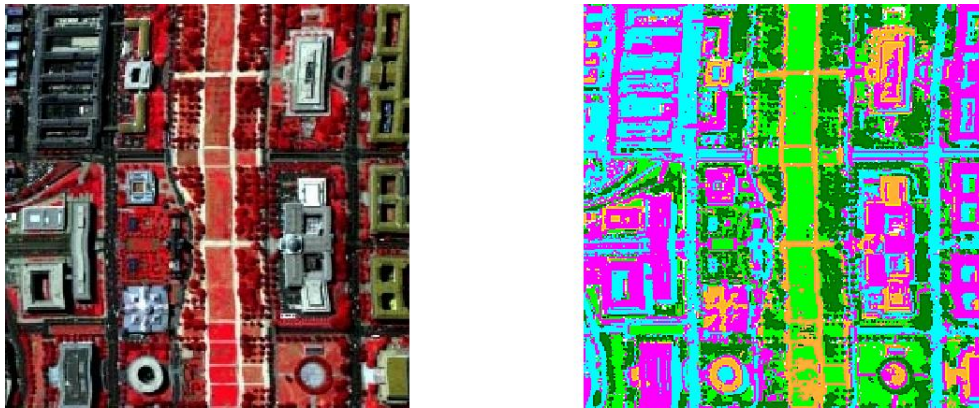


Figure 3. DC Mall: Comparison between original image and map produced by proposed method.

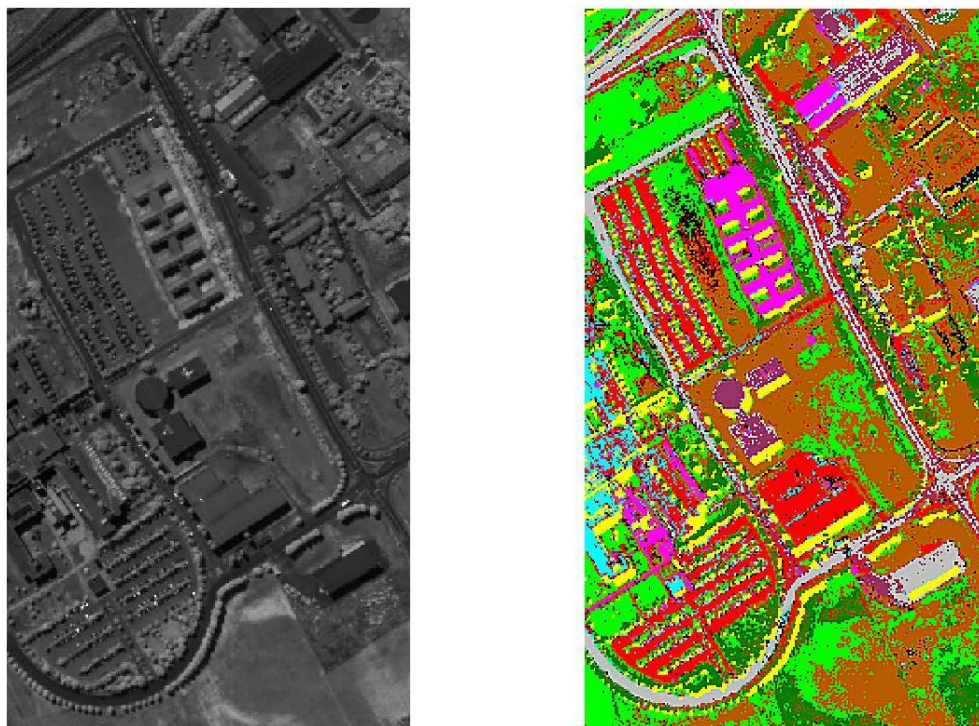


Figure 4. University: Comparison between original image and map produced by proposed method.

number of input nodes is determined by the dimension of the transform-based feature vector. The number of output nodes is equal to the number of classes in the image. The number of hidden layer nodes is set as equal to the square root of product of the number of input layer nodes and output layer nodes [23].

3.3. Training and Test Set

As limited numbers of labeled pixels are practically available because of the complexity and cost involved in sample collection, we kept the size of the training and test set small. The training and test pixels for both data sets are obtained with the help of the labeled field map available from the data source. In both cases, the training and test data sets are mutually exclusive and randomly selected.

3.4. Feature Extraction

Let $X = [X^1, X^2, \dots, X^S]$ where X^δ represents the matrix containing the hyperspectral signature of the training

pixels belonging to land cover class S identified in the image. Thus, X consists of the training samples of each class. The proposed one-dimensional adaptive lifting wavelet transform is applied to each signature kept row wise in matrix X up to the desired level. Thus, in case of the DC mall data set, the original hyperspectral signature of dimension $[1, 192]$ is converted into the wavelet-based feature vector of dimension $[1, 12]$ after decomposition up to level 4. For the Pavia University data, after two-level decomposition, the feature vector dimension gets reduced from $[1, 96]$ to $[1, 24]$. Since approximate signal obtained by the proposed ALS retains useful features, the approximate coefficients are used to construct wavelet-based feature vector.

3.5. Train ANN Classifier

Using the adaptive wavelet-based feature vectors, the single hidden layer feed forward neural network is trained by backpropagation algorithm. The network training is stopped after achieving termination criterion of either getting the desired test sample accuracy or upon reaching the maximum number of iterations. On completion of training, neural network is employed as a classifier to classify each image pixel signature into one of the land cover classes as follows.

3.6. ANN as Classifier

The trained ANN classifier is then used to classify every pixel's hyperspectral signature as follows.

- 1) Decompose original hyperspectral signature of a pixel up to the desired level using proposed adaptive lifting wavelet transform.
- 2) Use only the approximate coefficient as transform domain feature vector to feed into the trained network.
- 3) A given pixel is assigned the class of the output node having highest value.
- 4) Repeat these steps for all image pixels to generate the classification map.

4. Result Discussion

The performance is evaluated in terms of overall accuracy and kappa value, which are calculated as follows.

- 1) Overall accuracy (OA) is defined as ratio of the number of correctly classified samples to the total number of samples. It is computed from the confusion matrix as follows.

$$OA = \frac{\sum_{i=1}^C n_{ii}}{\sum_{i=1}^C \sum_{j=1}^C n_{ij}} \quad (6)$$

where n_{ij} is the element of the confusion matrix and denotes the number of samples of j^{th} ($j = 1, \dots, C$) class classified into i^{th} ($i = 1, \dots, C$) class. Here C represents the number of classes of the given data set.

- 2) Kappa coefficient (K) is computed from the confusion matrix. It is based on the difference between the actual agreement and the chance agreement (row and column totals) [24]. The value of Kappa coefficient lies in the range $[-1, +1]$. Better the classification, closer is the value of Kappa to $+1$. Mathematically, it is defined as,

$$K = \frac{n \sum_{i=1}^C n_{ii} - \sum_{i=1}^C n_{i+} n_{+i}}{n^2 - \sum_{i=1}^C n_{i+} n_{+i}} \quad (7)$$

where n denotes the total number of test samples and C denotes the number of classes of the given data set. Also n_{i+} denotes the sum of the elements if the i^{th} row and n_{+j} denotes the sum of the elements of column j .

The performance of proposed method is compared with first generation DWT-based feature extraction method. The results of the experiments on both data sets reported in **Table 1** are of over 10 simulations. From **Table 1**, it is observed that the proposed ALS-based feature extraction method gives best performance for both data sets. For first generation DWT-based feature extraction method, detail coefficients are used to construct the feature vector as they give better results than their approximate coefficients. For the proposed approach, approximate coefficients give the best result that detail one. This validates our algorithm because the proposed adaptive update step

retains the distinctive information in the approximate signal, not in detail. It is found that most of the detail coefficients obtained with ALS have very small value.

From **Figure 3**, it is observed that all classes identified in the original image are well classified and the obtained map has homogeneous regions. The map shows finer details and that similar class-Roof and Road are more accurately classified. The DWT-based approach fails to classify the finer details, and similar classes are not well separated. From the earlier studies on the same DC mall data set, it was found that the Nonparametric Weighted Feature Extraction (NWFE) algorithm by Kuo and Landgrebe gives 92% accuracy using Gaussian classifier with 12 features and 50 training samples of each class. The 2NN classifier with NWFE algorithm gives 87% accuracy [25]. **Table 2** shows the comparison of our result with the previous results. The proposed approach shows improvement in accuracy compared to previous studies.

For the University scene, it is observed from **Figure 4**, that all nine classes are well classified. It is clear that Meadows and Trees are more accurately classified using the proposed method, although the spectral reflectance of both classes is very similar. Likewise, the classes Asphalt and Bitumen are also well defined. The proposed approach also clearly classifies small objects such as Trees and Shadows. All regions with boarder shadows are very well defined. These results are compared with the results of previous studies that used the same data set. **Table 3** includes the results of previous algorithms: morphological-based classification using support vector machine (SVM), principal components and extended morphological profiles based classification by Plaza *et al.* [26]. It also includes results obtained with the pixel-wise classification, followed by majority voting within the watershed regions obtained by Tarabalka *et al.* [27]. The visual inspection of maps produced by the different algorithms mentioned in **Table 3** reveals that the proposed ALS can classify some regions more accurately such as meadows and bare soil and bricks on highway.

It is noted that the overall accuracy and kappa value are lower in the results of university scene compare to DC mall. This is due to the class distribution of University data set, even labeled training sample cannot represent the class distribution over all region. For both data sets, the overall classification accuracies for the classes that are represented by a few training samples are high. For instance, in the case of University scene, 99% accuracy was obtained for the Metal sheet class.

Table 1. Best overall accuracy, kappa value, mean and standard deviation for both data set.

Feature extraction	Features used	DC Mall				University scene			
		K	OA (%)	Mean	SD	K	OA (%)	Mean	SD
Proposed ALS	Approximate	0.96	97.2	95.43	2.35	0.81	83.4	82.8	0.64
First generation DWT									
Haar	Detail	0.88	90.1	90	0.15	0.66	69	65.2	2.49
Db4	Detail	0.93	95.23	91.42	3.76	0.65	67.7	58	3.7

Table 2. DC mall: Comparison with previous studies.

Algorithm	OA (%)	K
Proposed ALS	97.2	0.96
NWFE + Gaussian classifier	92.00	-
NWFE + 2NN	87	-

Table 3. University: Comparison with previous studies.

Algorithm	OA (%)	K
Proposed ALS	83.40	0.81
Morphology & SVM	81.01	0.75
PCA & EMP	85.22	0.80
SVM & majority voting	85.42	0.81

5. Conclusion

This paper presents the new adaptive update scheme to extract lower dimensional feature vector for the classification of hyperspectral signatures. The proposed approach preserves the distinctive features of hyperspectral signatures in the approximate signal, which lead to improvement in the classification result. It is clear that using the proposed approach, each class identified in both images is well separated. In fact, with the proposed approach, we are able to accurately distinguish the classes having similar spectral reflectance, for example, the classes such as Roof and Road in the case of DC mall and Meadows and Trees for the University of Pavia image. Another important aspect is that it gives better results even using a very small training set. This paper clearly demonstrates that our method is most robust for the classification of hyperspectral signatures with respect to overall accuracy, kappa value, and training data size.

Acknowledgements

The authors would like to thank David Landgrebe and Paolo Gamba for providing the data set.

References

- [1] Chang, C.I. (2003) *Hyperspectral Imaging: Techniques for Spectral Detection and Classification*. Kluwer, New York. <http://dx.doi.org/10.1007/978-1-4419-9170-6>
- [2] Donoho, D.L. (2000) *High-Dimensional Data Analysis: The Curses and Blessing of Dimensionality*. AMS Mathematical Challenges of the 21st Century.
- [3] Serpico, S.B. and Moser, G. (2007) Extraction of Spectral Channels From Hyperspectral Images for Classification Purposes. *IEEE Transactions on Geoscience and Remote Sensing*, **45**, 484-495. <http://dx.doi.org/10.1109/TGRS.2006.886177>
- [4] Mallat, S.G. (1989) A Theory for Multiresolution Signal Decomposition: The Wavelet Representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **11**, 674-693. <http://dx.doi.org/10.1109/34.192463>
- [5] Selesnick, I.W., Baraniuk, R.G. and Kingsbury, N.G. (2005) The Dual-Tree Complex Wavelet Transform. *IEEE Signal Processing Magazine*, **22**, 123-151. <http://dx.doi.org/10.1109/MSP.2005.1550194>
- [6] Hsu, P.-H., Tseng, Y.-H. and Gong, P. (2006) Spectral Feature Extraction of Hyperspectral Images Using Wavelet Transform. *Journal of Photogrammetry and Remote Sensing*, **11**, 93-109.
- [7] Bruce, L.M., Koger, C.H. and Li, J. (2002) Dimensionality Reduction of Hyperspectral Data Using Discrete Wavelet Transform Feature Extraction. *IEEE Transactions on Geoscience and Remote Sensing*, **40**, 2331-2338.
- [8] Choi, H., Romberg, J., Baraniuk, R. and Kingsbury, N. (2000) Hidden Markov Tree Modeling of Complex Wavelet Transforms. *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, **1**, 133-136.
- [9] Sweldens, W. (1995) The Lifting Scheme: A New Philosophy in Biorthogonal Wavelet Constructions. In: *Proceedings of SPIE 2569: Wavelet Applications in Signal and Image Processing III*, SPIE Press, San Diego, 68-79.
- [10] Claypoole, R.L., Baraniuk, R.G. and Nowak, R.D. (1998) Adaptive Wavelet Transforms via Lifting. *Proceedings of the 1998 IEEE International Conference on Acoustics, Speech and Signal Processing*, **3**, 1513-1516. <http://dx.doi.org/10.1109/icassp.1998.681737>
- [11] Claypoole, R.L., Davis, G.M., Sweldens, W. and Baraniuk, R.G. (2003) Nonlinear Wavelet Transforms for Image Coding via Lifting. *IEEE Transactions on Image Processing*, **12**, 1449-1459. <http://dx.doi.org/10.1109/TIP.2003.817237>
- [12] Piella, G. and Heijmans, H.J.A.M. (2002) Adaptive Lifting Schemes with Perfect Reconstruction. *IEEE Transactions on Signal Processing*, **50**, 1620-1630. <http://dx.doi.org/10.1109/TSP.2002.1011203>
- [13] Heijmas, H.J.A.M., Pesquet-Popescu, B. and Piella, G. (2005) Building Nonredundant Adaptive Wavelets by Update Lifting. *Applied and Computational Harmonic Analysis*, **18**, 252-281. <http://dx.doi.org/10.1016/j.acha.2004.11.006>
- [14] Piella, G. and Pesquet-Popescu, B. (2007) A Three-Step Nonlinear Lifting Scheme for Lossless Image Compression. *Proceedings of the IEEE International Conference on Image Processing*, San Antonio, 16 September-19 October 2007, 453-456.
- [15] Tomic, M. and Sersic, D. (2013) Point-Wise Adaptive Wavelet Transform for Signal Denoising. *INFORMATICA*, **24**, 637-656.
- [16] Subasi, A. and Erçelebi, E. (2005) Classification of EEG Signals Using Neural Network and Logistic Regression. *Computer Methods and Programs in Biomedicine*, **78**, 87-99. <http://dx.doi.org/10.1016/j.cmpb.2004.10.009>
- [17] Ding, X.W. and Huang, W.G. (2007) SAR Image De-Noiseing by Wavelet Transform Based on Lifting Scheme. *Pro-*

- ceedings of SPIE—The International Society for Optical Engineering*, 67-87. <http://dx.doi.org/10.1117/12.748762>
- [18] Gouze, A., Antonini, M., Barlaud, M. and Macq, B. (2004) Design of Signal-Adapted Multidimensional Lifting Scheme for Lossy Coding. *IEEE Transactions on Image Processing*, **13**, 1589-1603.
- [19] Benazza-Benyahia, A., Pesquet, J.-C., Hattay, J. and Masmoudi, H. (2007) Block-Based Adaptive Vector Lifting Schemes for Multichannel Image Coding. *EURASIP Journal on Image and Video Processing*, **2007**, Article ID: 013421.
- [20] Landgrebe, D.A. (2003) *Signal Theory Methods in Multispectral Remote Sensing*. John Wiley and Sons, New York. <http://dx.doi.org/10.1002/0471723800>
- [21] <https://engineering.purdue.edu/biehl/MultiSpec/hyperspectral>
- [22] Gamba, P. University of Pavia. <http://eecs.unipv.it/teaching-staff/paolo-gamba/>
- [23] Haykin, S. (1997) *Neural Networks: A Comprehensive Foundation*. Prentice Hall, England.
- [24] Richards, J.A. and Jia, X. (1999) *Remote Sensing Digital Image Analysis: An Introduction*. Springer, New York. <http://dx.doi.org/10.1007/978-3-662-03978-6>
- [25] Kuo, B.-C. and Landgrebe, D.A. (2004) Hyperspectral Data Classification Using Nonparametric Weighted Feature Extraction. *IEEE Transactions on Geoscience and Remote Sensing*, **42**, 1096-1105. <http://dx.doi.org/10.1109/TGRS.2004.825578>
- [26] Plaza, A., Benediktsson, J.A., Boardman, J.W., Braziled, J., Bruzzonee, L., Camps-Valls, G., *et al.* (2009) Recent Advances in Techniques for Hyperspectral Image Processing. *Remote Sensing of Environment*, **113**, S110-S122.
- [27] Tarabalka, Y., Chanussot, J. and Benediktsson, J.A. (2010) Segmentation and Classification of Hyperspectral Images Using Watershed Transformation. *Pattern Recognition*, **43**, 2367-2379. <http://dx.doi.org/10.1016/j.patcog.2010.01.016>