

AN EXTENDED SPECTRAL–SPATIAL CLASSIFICATION APPROACH FOR HYPERSPECTRAL DATA

D. Akbari^a

^a Surveying and Geomatics Engineering Department, College of Engineering, University of Zabol, Zabol, Iran -
davoodakbari@uoz.ac.ir

KEY WORDS: Hyperspectral image, Spectral-spatial classification, Subspace extraction, Support Vector Machines, Minimum Spanning Forest, Segmentation map

ABSTRACT:

In this paper an extended classification approach for hyperspectral imagery based on both spectral and spatial information is proposed. The spatial information is obtained by an enhanced marker-based minimum spanning forest (MSF) algorithm. Three different methods of dimension reduction are first used to obtain the subspace of hyperspectral data: (1) unsupervised feature extraction methods including principal component analysis (PCA), independent component analysis (ICA), and minimum noise fraction (MNF); (2) supervised feature extraction including decision boundary feature extraction (DBFE), discriminate analysis feature extraction (DAFE), and nonparametric weighted feature extraction (NWFE); (3) genetic algorithm (GA). The spectral features obtained are then fed into the enhanced marker-based MSF classification algorithm. In the enhanced MSF algorithm, the markers are extracted from the classification maps obtained by both SVM and watershed segmentation algorithm. To evaluate the proposed approach, the Pavia University hyperspectral data is tested. Experimental results show that the proposed approach using GA achieves an approximately 8% overall accuracy higher than the original MSF-based algorithm.

1. INTRODUCTION

Hyperspectral imaging concerns the measurement, analysis, and interpretation of spectral imagery acquired from either a given scene or a specific object by satellite, airborne, terrestrial, or laboratory sensor, over visible, infrared and sometime thermal spectral regions of electromagnetic spectrum (Shippert, 2004). Recent technological improvements in spatial, spectral, and radiometric characteristics of spectrometer imagers beget the need of developing new methods for land cover mapping. There are two major approaches for classification of hyperspectral images: the spectral or pixel-based and the spectral-spatial approaches. While the pixel-based techniques, such as the classic Maximum Likelihood or Support Vector Machines (SVM) classifiers, primarily emphasize the independence of pixels, the spectral-spatial frameworks such as Geographic Object-Based Image Analysis (GEOBIA) (Blaschke et al., 2014) or Minimum Spanning Forest (MSF) (Tarabalka et al., 2010a) classifiers employ both the spectral characteristics and the spatial context of the pixels. Many researchers have demonstrated that the use of spectral-spatial information improves the classification results, compared to the use of spectral data alone, in hyperspectral imagery (Plaza et al., 2009; Li et al., 2010; Fauvel et al., 2012; Heras et al., 2014; Xu et al., 2014).

Segmentation techniques are powerful means for defining the spatial dependences among the pixels and for finding the homogeneous regions in an image (Gonzalez and Woods, 2002; Chen et al., 2012). The advantages of using segmentation for distinguishing spatial structures from one another are also discussed in (Tarabalka et al., 2009; Tarabalka et al., 2010; Bitam and Ameur, 2013). An alternative way in order to improve the accuracy of segmentation is performing a marker-based technique (Soille, 2003; Tarabalka et al., 2010). The idea behind this approach is selecting of one or several pixels for every spatial object as the seed or a marker of the corresponding region. Marker-based segmentation considerably decreases the over-segmentation. As a result, it leads to more reliable accuracies.

For classification of the high-dimensional hyperspectral images, there is a major problem. The problem is due to the high number of spectral channels and the relatively small number of labels samples. Many algorithms have been reported to be effective in reducing the dimensions of input space and achieving better performance, such as principal component analysis (PCA) (Srivastava and Liu, 2005; Saegusa et al., 2004), independent component analysis (ICA) (Zheng et al., 2006), maximum noise fraction (MNF) (Green et al., 1988) unsupervised feature extraction methods, decision boundary feature extraction (DBFE) (Lee and Landgrebe, 1993), discriminate analysis feature extraction (DAFE) (Landgrebe, 2003), and nonparametric weighted feature extraction (NWFE) (Kuo and Landgrebe, 2004) supervised feature extraction methods.

In this paper, we propose an extended spectral-spatial classification approach based on subspace analysis of hyperspectral remote sensing data. In the proposed approach, three different methods are employed to extract the optimal hyperspectral features, including unsupervised and supervised feature extraction, and the genetic algorithm (GA). Afterwards, the enhanced marker-based MSF (MMSF) spectral-spatial algorithm is used to classify the optimal features. In the enhanced MMSF algorithm to select markers the pixels related to a given class with the largest population are kept for each region of the segmentation map. Finally, the most reliable labelled pixels are selected among the existing pixels for each region as the markers. Therefore, it benefits from a segmentation algorithm to integrate the spatial information into the marker selection process.

2. LITERATURE REVIEW

Automatic marker selection has been previously used in the literature mostly for the greyscale and colour images. Markers are often chosen by searching the flat zones (i.e. the connected components of pixels with a constant gray level value) or the zones of homogeneous texture (Soille, 2003). Noyel et al. (2007;

2008) performed classification of hyperspectral image using different methods, such as Clara (Kaufman and Rousseeuw, 1990) and linear discriminate analysis (Duda et al., 2001) and then filtered the classification maps, using mathematical morphology operators, for selecting large spatial regions as markers. In (Tarabalka et al., 2010a) an efficient approach is proposed for spectral-spatial classification using the MSF grown from automatically selected markers. It uses a pixel-wise SVM classification, in order to select pixels with the highest probability estimate to each class, as markers. In this approach, a connected components labelling is, first, applied on the classification map. Then, the markers are considered to be $p\%$ of the pixels with the highest probability estimate for large regions, and pixels with an estimated probability higher than a pre-defined threshold for small regions. The disadvantage of this approach is that it does not employ the spatial or neighbouring information in marker selection process.

By summarizing the exiting literature about subspace extraction, it can be found that it always focuses on the pixel-based classification, without considering the spatial relationship of neighbouring pixels. Recent studies show that the exploitation of spatial information is necessary for classification of hyperspectral imagery, but few such approaches have been proposed (Huang and Zhang, 2009), which is partly due to the high dimensionality of the data and the spectral and spatial heterogeneity of remote sensing images (Duarte-Carvajalino, 2008).

3. THE PROPOSED APPROACH

The flowchart of the proposed spectral-spatial classification approach is presented in Figure 1.

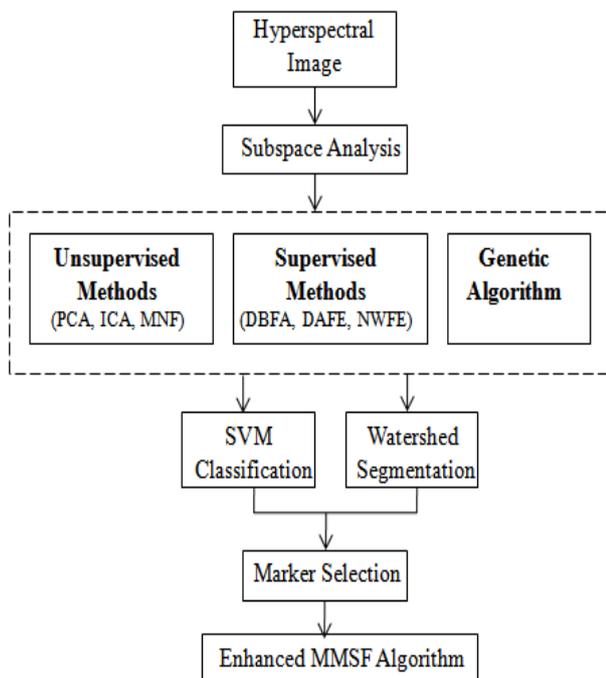


Figure 1. Schema of the proposed approach.

As shown, the proposed approach consists of two blocks:

(1) Subspace extraction for pre-processing: it aims to reduce the dimensionality and extract the spectral subspace from hyperspectral data. For this purpose, three different techniques are employed, including unsupervised and supervised feature extraction, and GA. In the unsupervised feature extraction method, we use PCA, ICA and MNF algorithms. The DBFE, DAFE and NWFE algorithms are used in the supervised feature extraction method. The GA is a general adaptive optimization search method based on a direct analogy to Darwinian natural selection and genetics in biological systems (Huang and Wang, 2006). It starts from an initial population which is composed of a set of possible solutions called individuals (chromosomes), and then evaluates the quality of each individual based on a fitness function. We use the accuracy parameters of SVM classification obtained from the training samples subset as the fitness function. The fitter solutions have a better chance to survive or reproduce in the next generations. The population during consecutive generations evolves to be fitter in the problem's conditions. Selection, crossover, and mutation are the main GA's operators to reproduce future generations. The evolutionary process will not stop until termination conditions satisfy (Zhuo and Zheng, 2008).

(2) The enhanced MMSF spectral-spatial classification of the spectral subspace: the SVM and the watershed segmentation algorithm are first used, in parallel, for classification and for segmentation of the spectral subspace extracted, respectively. Watershed segmentation is an efficient morphological algorithm for image segmentation. It combines both region growing and edge detection techniques (Vincent and Soille, 1991). Afterwards, all the pixels related to the class with the largest population are kept for each region of segmentation map (see Figure 2). Lastly, the most reliable labelled pixels are selected among the existing pixels for each region as markers. The markers are then used to build the MSF.

4. EXPERIMENTAL RESULTS AND DISCUSSIONS

The Pavia University hyperspectral image is used for our experiments. It is collected by ROSIS-03 sensor over Pavia city. Table 1 describes the main characteristics of this dataset.

Dataset	Pavia University
Sensor	ROSI3-03
Spectral range (um)	0.43-0.86
Spatial coverage (pixel)	610×340
Spatial resolution (m)	1.3
Number of bands	103
Number of classes	9

Table 1. The main characteristics of the dataset used.

4.1 Experiment results

In this study, to perform PCA, NWFE and DBFE algorithms, we kept the spectral subspace with the total variance more than 90%. The MNF eigenimages with near-unity eigenvalues can be viewed as noise-dominated features and hence removed. It should be noted that the maximum $N-1$ dimensional subspace is available for DAFE since the maximum rank of $\sum B$, a between-class scatter matrix, is $N-1$ for an N -class classification problem (Landgrebe, 2003). Table 2 presents the value of parameters used in GA.

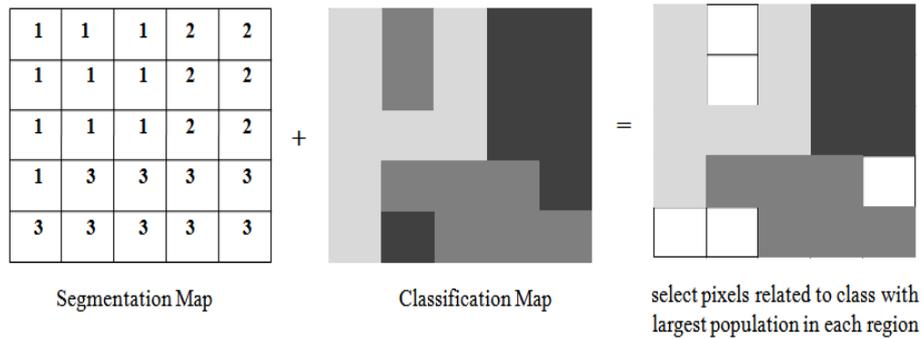


Figure 2. An example of the interference segmentation map in SVM classification map.

Parameters	data
Population	100
Crossover probability	80%
Mutation probability	0.9%
K-tournament	2
K-elitism	2

Table 2. The GA's Parameters for dataset used.

Classes	No. of training set	No. of test set
Asphalt	548	6304
Meadows	540	18146
Gravel	392	1815
Trees	524	2912
Metal sheets	265	1113
Bare Soil	532	4572
Bitumen	375	981
Bricks	514	3364
Shadows	231	795

Table 3. The training and test samples.

In experiments, the Gaussian radial basis function (RBF) kernel is used for the SVM classifier (Camps-Valls and Bruzzone, 2005). The RBF kernel's parameters, i.e. C and γ , are chosen by a five-fold cross validation. In order to create a map of markers in enhanced MMSF algorithm, as mentioned in section 3, first, for each region of segmentation map, all the pixels related to the class with the largest population are kept. Then, if it contains more than 40 pixels, 9% of its pixels with the highest estimated probability are selected as the marker. Otherwise, the region marker is formed by the pixels with an estimated probability higher than a threshold τ . The threshold τ is equal to the lowest probability within the highest 6% of the probabilities for the whole image. In the next step, the image pixels are grouped into the MSF using the spectral angle dissimilarity measure built from the selected markers (Van der Meer, 2006). In order to compare the results of the proposed approach we have implemented the original and enhanced MMSF algorithms on all image bands. In the original MMSF algorithm the labeling of connected components is performed using the eight-neighborhood connectivity. For each connected component, if it contains more than 20 pixels, 5% of its pixels with the highest estimated probability are selected as the marker for this component. Otherwise, the region marker is formed by the pixels with an estimated probability higher than 2%.

The accuracies of the classification maps are generally assessed by the overall accuracy (OA), the Kappa coefficient of agreement (κ), and the class-specific producer's accuracy (PA). The OA is the percentage of correctly classified pixels, the κ is the percentage of agreement corrected by the amount of agreement that can be expected due to chance alone, and the PA is the percentage of correctly classified samples for a given class.

4.2 Discussions

In the Pavia University dataset, a total of 3921 and 40002 pixels were available as training and test data, respectively (see Table 3).

Figure 3 shows the color composite image, reference data and the classification maps of Original-MMSF, Enhanced-MMSF and the proposed approach. In this figure, the classification map obtained by GA-MMSF contains much more homogeneous regions when compared with the maps obtained by other algorithms (see Figure 3(k)).

Table 4 presents the number of features and accuracy values estimated for different methods. From the results, it can be found that all subspaces gave higher accuracies than the Original-MMSF that uses 104 channels. Also, in the unsupervised and supervised feature extraction methods, PCA and DAFE algorithms provide the best results, respectively. However, as can be seen from Table 4, the best results are achieved using GA subspace features. The kappa coefficient of GA-MMSF algorithm is approximately 9% higher than the Original-MMSF algorithm.

In Table 4, all the class-specific producer's accuracy rates for the proposed GA-MMSF algorithm are higher than Original-MMSF approach.

5. CONCLUSION

In this study, an improved spectral-spatial classification approach for hyperspectral images based on the subspace analysis techniques has been proposed. The subspace analysis techniques are used to reduce the computational cost, since the spectral-spatial classification is time-consuming and unacceptable for hyperspectral data with hundreds of channels. On the other hand, subspace analysis is able to reduce the information redundancy in hyperspectral data as the huge spectral channels are highly correlated. Therefore, we proposed to integrate the subspace analysis and spectral-spatial classification for hyperspectral image interpretation. In the proposed approach, the dimension reduction of hyperspectral images is accomplished using three different approaches

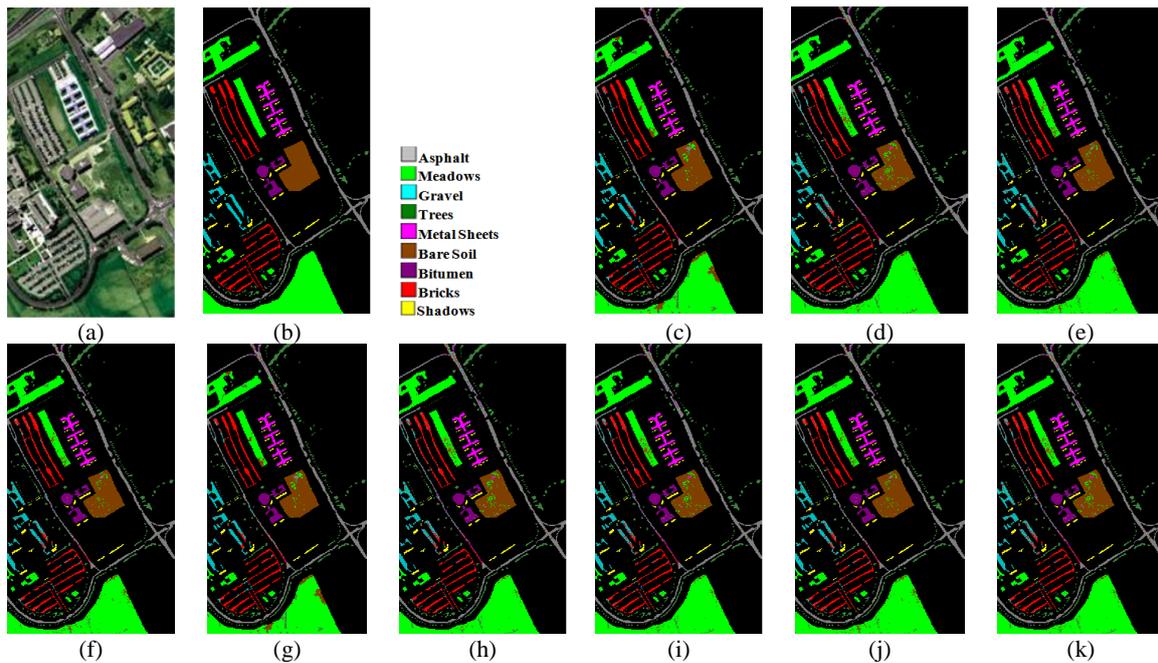


Figure 3. (a) Colour composite image, (b) Reference map, (c) Original-MMSF classification map, (d) Enhanced-MMSF classification map, (e) PCA-MMSF classification map, (f) ICA-MMSF classification map, (g) MNF-MMSF classification map, (h) DBFE-MMSF classification map, (i) DAFE-MMSF classification map, (j) NWFE-MMSF classification map, (k) GA-MMSF classification map.

Methods	Original-MMSF	Enhanced-MMSF	PCA-MMSF	ICA-MMSF	MNF-MMSF	DBFE-MMSF	DAFE-MMSF	NWFE-MMSF	GA-MMSF
No. of Features	104	104	32	17	17	15	8	15	54
OA (%)	78.6	82.0	82.2	80.4	79.4	83.2	83.9	81.1	86.7
κ (%)	73.3	77.7	78.0	77.3	76.1	80.1	80.8	77.1	82.8
Asphalt	95.8	95.9	93.3	92.0	91.6	96.4	95.5	93.1	96.3
Meadows	62.2	75.5	61.6	57.6	57.5	73.0	73.1	74.3	80.1
Gravel	94.7	94.4	94.2	92.3	92.1	95.2	94.9	91.2	96.0
Trees	96.3	97.8	96.5	97.1	97.7	96.8	97.2	92.8	98.3
Metal Sheets	88.0	88.7	88.1	88.0	86.1	88.4	88.1	87.3	88.8
Bare Soil	85.1	88.0	86.3	84.7	81.0	84.9	85.2	84.5	85.4
Bitumen	97.5	95.5	97.4	91.4	92.2	96.2	97.2	96.5	95.7
Bricks	96.7	97.0	96.3	96.1	95.2	97.1	96.6	96.0	96.9
Shadows	56.3	62.9	54.8	52.7	65.2	63.0	65.9	57.2	65.9

Table 4. The number of features and accuracy values obtained.

including unsupervised and supervised feature extraction methods and the GA. The enhanced MMSF spectral-spatial algorithm is then used to classify extracted features. In the proposed MMSF algorithm, the corresponding pixels of each class with the largest population for each region of watershed segmentation map are kept. Then, the most reliable labelled pixels are selected as markers and used to build the MSF. Experimental results show that the subspace images are effective in extracting spectral information from the hyperspectral data.

The proposed methodology was able to take advantage of spectral and spatial information simultaneously for accurate classification of hyperspectral image. Further work is needed to improve the proposed approach. It is, for example, necessary to take advantage of the available data in order to automate the entire classification process.

ACKNOWLEDGEMENTS

The authors would like to thank Prof. P. Gamba from the University of Pavia, Italy, for providing the hyperspectral data.

REFERENCES

- Bitam, A., and Ameer, S., 2013. A local-spectral fuzzy segmentation for MSG multispectral images. *International Journal of Remote Sensing*, 34: 8360–8372.
- Blaschke, T., Hay, G.J., Kelly, M., Lang, S., Hofmann, P., Addink, E., and Tiede, D., 2014. Geographic Object-Based Image Analysis – Towards a new paradigm. *ISPRS Journal of Photogrammetry and Remote Sensing*, 87: 180–191.
- Camps-Valls, G., and Bruzzone, L., 2005. Kernel-based methods for hyperspectral image classification. *IEEE Translation Geoscience Remote Sensing*, 43: 1351–1362.
- Chen, Z., Wang, G., and Liu, J., 2012. A modified object-oriented classification algorithm and its application in high-resolution remote-sensing imagery. *International Journal of Remote Sensing*, 33: 3048–3062.

- Duarte-Carvajalino, J.M., Sapiro, G., Velez-Reyes, M., and Castillo, P.E., 2008. Multiscale representation and segmentation of hyperspectral imagery using geometric partial differential equations and algebraic multigrid methods. *IEEE Transactions on Geoscience and Remote Sensing*, 46(8), 2418–2434.
- Duda, R.O., Hart, P.E., and Stork, D.G., 2001. *Pattern Classification*, New York: Wiley.
- Fauvel, M., Chanussot, J., and Benediktsson, J.A., 2012. A spatial–spectral kernel-based approach for the classification of remote sensing images. *Pattern Recognition*, 45: 381–392.
- Gonzalez, R., and Woods, R., 2002. *Digital Image Processing*. 2nd ed. Englewood Cliffs, NJ: Prentice-Hall.
- Green, A.A., Berman, M., Switzer, P., and Craig, M.D. 1988. A transformation for ordering multispectral data in terms of image quality with implications for noise removal. *IEEE Translation Geoscience Remote Sensing*, vol. 26, no. 1, pp. 65.
- Heras, D.B., Argüello, F., and Quesada-Barriuso, P., 2014. Exploring ELM-based spatial–spectral classification of hyperspectral images. *International Journal of Remote Sensing*, 35(2): 401–423.
- Huang, C.-L., and Wang, C.-J., 2006. A GA-based feature selection and parameter optimization for support vector machines. *Expert Systems with Application*, 231–240.
- Huang, X., and Zhang, L., 2009. A comparative study of spatial approaches for urban mapping using hyperspectral ROSIS images over Pavia city, northern of Italy. *International Journal of Remote Sensing*, 30(12), 3205–3221.
- Kaufman, L., and Rousseeuw, P.J., 1990. *Finding Groups in Data, An Introduction to Cluster Analysis*.- New York: Wiley.
- Kuo, B., and Landgrebe, D.A., 2004. Nonparametric weighted feature extraction for classification. *IEEE Transactions on Geoscience and Remote Sensing*, 42 (5), 1096–1105.
- Landgrebe, D.A., 2003. *Signal Theory Methods in Multispectral Remote Sensing*. Wiley, Hoboken, NJ.
- Lee, C., and Landgrebe, D.A., 1993. Feature extraction based on decision boundaries. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15 (4), 388–400.
- Li, H., Gu, H., Han, Y., and Yang, J., 2010. Object-oriented classification of high-resolution remote sensing imagery based on an improved colour structure code and a support vector machine. *International Journal of Remote Sensing*, 31: 1453–1470.
- Noyel, G., 2008. *Filtrage, Réduction de Dimension, Classification et Segmentation Morphologique Hyperspectrale*, Ph.D. dissertation, Ctr. Mathematical Morphology, Paris Sch. Mines, Paris, France.
- Noyel, G., Angulo, J., and Jeulin, D., 2007. Morphological segmentation of hyperspectral images. *Image Anal Stereol*, 26: 101–109.
- Plaza, A., Benediktsson, J.A., Boardman, J.W., Brazile, J., Bruzzone, L., Camps-Valls, G., and Trianni, G., 2009. Recent advances in techniques for hyperspectral image processing. *Remote Sensing of Environment*, 113: S110–S122.
- Saegusa, R., Sakano, H., and Hashimoto, S., 2004. Nonlinear principal component analysis to preserve the order of principal components. *Neurocomputing*, 61 57–70.
- Shippert, P., 2004. Why Use Hyperspectral Imagery?. *Photogrammetric Engineering and Remote Sensing*, 377–380.
- Soille, P., 2003. *Morphological Image Analysis*. 2nd ed. Berlin, Germany: Springer-Verlag.
- Srivastava, A., and Liu, X., 2005. Tools for application-driven linear dimension reduction. *Neurocomputing*, 67 136–160.
- Tarabalka, Y., Benediktsson, J.A., and Chanussot, J., 2009. Spectral-spatial classification of hyperspectral imagery based on partitioning clustering techniques. *IEEE Translation Geoscience Remote Sensing*, 47(9): 2973–2987.
- Tarabalka, Y., Benediktsson, J.A., Chanussot, J., and Tilton, J.C., 2010. Multiple spectral-spatial classification approach for hyperspectral data. *IEEE Translation Geoscience Remote Sensing*, 48(11): 4122–4132.
- Tarabalka, Y., Chanussot, J., and Benediktsson, J.A., 2010a. Segmentation and classification of hyperspectral images using minimum spanning forest grown from automatically selected markers. *IEEE Translation System, Man, Cybern. B, Cybern.*, 40(5): 1267–1279.
- Van Der Meer, F., 2006. The effectiveness of spectral similarity measures for the analysis of hyperspectral imagery. *International Journal Apply Earth Observation Geoinformation*, 8: 3–17.
- Vincent, L., and Soille, P., 1991. Watersheds in digital spaces: an efficient algorithm based on immersion simulations. *IEEE Translation Pattern Anal. Mach. Intell.*, 583–598.
- Xu, J., Hang, R., and Liu, Q., 2014. Patch-based active learning (PTAL) for spectral-spatial classification on hyperspectral data. *International Journal of Remote Sensing*, 35(5): 1846–1875.
- Zheng, C., Huang, D., and Shang, L., 2006. Feature selection in independent component subspace for microarray data classification. *Neurocomputing*, 69 2407–2410.
- Zhuo, L., and Zheng, J., 2008. A Genetic Algorithm Based Wrapper Feature Selection Method for Classification of Hyperspectral Image Using Support Vector Machine. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 397–402.