

## A structural similarity-based label-smoothing algorithm for the post-processing of land-cover classification

Qikai Lu, Xin Huang, Tingting Liu & Liangpei Zhang

To cite this article: Qikai Lu, Xin Huang, Tingting Liu & Liangpei Zhang (2016) A structural similarity-based label-smoothing algorithm for the post-processing of land-cover classification, Remote Sensing Letters, 7:5, 437-445, DOI: [10.1080/2150704X.2016.1149252](https://doi.org/10.1080/2150704X.2016.1149252)

To link to this article: <http://dx.doi.org/10.1080/2150704X.2016.1149252>



Published online: 18 Feb 2016.



Submit your article to this journal [↗](#)



Article views: 36



View related articles [↗](#)



View Crossmark data [↗](#)

## A structural similarity-based label-smoothing algorithm for the post-processing of land-cover classification

Qikai Lu<sup>a,b</sup>, Xin Huang<sup>a,c</sup>, Tingting Liu<sup>d</sup> and Liangpei Zhang<sup>a</sup>

<sup>a</sup>State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University, Wuhan, China; <sup>b</sup>Electronic Information School, Wuhan University, Wuhan, China; <sup>c</sup>School of Remote Sensing and Information Engineering, Wuhan University, Wuhan, China; <sup>d</sup>Chinese Antarctic Center of Surveying and Mapping, Wuhan University, Wuhan, China

### ABSTRACT

Post-processing is able to achieve a satisfactory classification performance with a low cost and simple assumption, making it widely used in the refinement of classification maps. In this study, a novel structural similarity-based label-smoothing algorithm is developed for the post-processing of land-cover classification. Inspired by the non-local (NL) means algorithm, the proposed algorithm assigns different voting weights to the neighbouring pixels for the identification of the central pixel. Here, the voting weight of a specific neighbouring pixel depends on its structural similarity to the central pixel. In this paper, two measurements are proposed to evaluate the similarity between pixels: (1) a consistency criterion; and (2) a histogram similarity criterion. The proposed algorithm was tested on three remote-sensing images. The experimental results confirm that the proposed algorithm reduces the classification noise and preserves the detail and structural information at the same time. Compared to the traditional post-processing approaches (e.g., majority voting), the proposed algorithm exhibits a more satisfactory performance.

### ARTICLE HISTORY

Received 2 December 2015  
Accepted 25 January 2016

## 1. Introduction

In recent years, remote sensing has been proved to be an effective approach for Earth observation. However, with the rapid development of geospatial technologies, large amounts of remotely sensed data with growing spectral, spatial and temporal resolutions are being obtained by a number of sensors. In this context, manual interpretation cannot easily meet the needs of practical applications, due to its low efficiency and high cost. Classification, i.e., converting remotely sensed data into a semantically labelled image, has therefore become one of the widely studied methods for exploiting the data information. Nevertheless, owing to the complexity of the class distribution in the spectral domain, spectral-based classification is always subject to salt-and-pepper noise, resulting in an inaccurate thematic map (Myint et al. 2011).

With the increase in the spatial resolution, more contextual and structural information can be provided by remote-sensing images. To exploit the rich spatial information

contained in the images, several state-of-the-art algorithms, e.g., morphological profiles (Ghamisi, Mura, and Benediktsson 2015), the multiscale cluster histogram (MCH) (Lu, Huang, and Zhang 2014) and object-based image analysis (OBIA) (Liu and Xia 2010), have been developed for spectral-spatial classification. The approaches based on spatial feature extraction can be viewed as classification pre-processing. However, these approaches always refer to a lot of parameters, leading to a high feature dimension and a large memory requirement. Therefore, as an alternative approach, classification post-processing has been proposed. According to the fact that adjacent pixels are more likely to belong to the same class, classification post-processing can be used and regarded as a refinement of the labelling (Schindler 2012). Recently, Huang et al. (2014) reviewed and compared several classification post-processing methods, and confirmed that the post-processing approaches are effective in improving land-cover classification. Specifically, the majority voting (MV) filter, which assigns the majority class in the neighbourhood to the central pixel, is the most widely used post-processing approach. Furthermore, the distance-weighted majority voting (DWV) filter assigns different weights to the neighbouring pixels around the central pixel, based on the spatial distance. In this approach, a higher weight is given to a pixel that is closer to the centre in a window. Although the filtering-based post-processing strategy is effective in enhancing the classification performance, few studies have investigated this approach in depth.

In this context, a novel approach inspired by the non-local (NL) means algorithm is proposed to refine the classification map. The NL-means algorithm was first developed for image denoising (Buades, Coll, and Morel 2005). It assumes that the structural similarity between two pixels depends on the similarity between the geometrical configurations in their whole neighbourhoods (Ma, Zhao, and Yuille 2016). A pixel is assigned a higher weight if its neighbourhood is more similar to that of the central pixel. Accordingly, the proposed method utilizes the structural similarity between pixels to determine the weight of each pixel in the post-processing process. In this study, two structural similarity-based weighting schemes are proposed. The first approach refers to the label vector that represents a squared neighbourhood of a pixel. The other is based on the class frequencies within the local window. In the experiments undertaken in this study, a series of remote-sensing images acquired by different sensors were employed to validate the effectiveness of the proposed approach.

## 2. Methodology

### 2.1. Object-based classification post-processing method

The object-based voting (OBV) approach aims to refine the initial classification result by considering the boundary derived from the image segmentation (Huang et al. 2014). The OBV algorithm can be described as

$$C(x_i) = \arg \max_k \left( \sum_{x_j \in S} \text{eq}(B(x_j), c_k) \right) \quad (1)$$

$$\text{eq}(a, b) = \begin{cases} 1 & a = b \\ 0 & a \neq b \end{cases} \quad (2)$$

where  $x_i$  is a pixel contained in image and  $S$  is the segment which  $x_i$  belongs to.  $B$  and  $C$  are the initial classification and the post-processing result, respectively, and  $c_1, c_2, \dots, c_m$  represent the  $m$  class labels. The function  $\text{eq}()$  is used to determine whether the variations  $a$  and  $b$  have the same value. It should be noted that the process is carried out with the labels from the initial classification, rather than based on the refined labels.

## 2.2. Filtering-based classification post-processing method

Filtering is the most commonly used classification post-processing approach. Based on a sliding window, the class information from the neighbouring pixels is utilized for the identification of the test pixel. Specifically, the MV and DWV filters are the two representative approaches.

(1) Majority voting filter: the MV filter is effective in erasing small and isolated objects. Its principle is to reclassify each pixel as the predominant class occurring in its neighbourhood, which can be expressed as

$$C(x_i) = \arg \max_k \left( \sum_{x_j \in W} \text{eq}(B(x_j), c_k) \right) \quad (3)$$

where  $W$  represents the window centred at pixel  $x_i$ . Nevertheless, with an increase in the window size, more pixels are considered in the voting, which always leads to an over-smooth and boundary blurring effect.

(2) Distance-weighted majority voting filter: the DWV filter is an enhanced majority filter, which considers the spatial position of pixels in the image. In MV, all the pixels within the window are assigned the same weights, regardless of their distance to the window centre. However, in the DWV strategy, the weight of a specific neighbouring pixel is related to its distance to the centre. The following equation shows the DWV method with a Gaussian distance function:

$$C(x_i) = \arg \max_k \left( \sum_{x_j \in W} \text{eq}(B(x_j), c_k) G_\sigma(\|x_i - x_j\|) \right) \quad (4)$$

$$G_\sigma(x) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2}{2\sigma^2}} \quad (5)$$

where  $G_\sigma$  is a Gaussian function with standard deviation  $\sigma$ , and  $\|x_i - x_j\|$  denotes the spatial distance between  $x_i$  and  $x_j$ . In this way, a large weight is assigned to a pixel which is close to the centre, while a small weight is given to a distant pixel.

## 2.3. Proposed method

Although MV and DWV are effective in smoothing the label image, they tend to blur the boundaries to some extent since the same weights are assigned to the pixels with the same offset to the central pixel. In this study, a novel classification post-processing

approach is proposed. Inspired by the NL-means algorithm (Buades, Coll, and Morel 2005), the structural information of the labelled image is used to smooth the homogeneous areas, but at the same time preserve the object boundaries.

The proposed approach compares the geometrical configurations in the pixels' neighbourhoods by considering the spatial distribution of the class labels within a neighbourhood for each pixel. It assigns larger weights to the pixels whose neighbourhoods are similar to that of the central pixel. Accordingly, in this study, the structural similarity-based label-smoothing algorithm is proposed as

$$C(x_i) = \arg \max_k \left( \sum_{x_j \in W} \text{eq}(B(x_j), c_k) \cdot \omega(x_i, x_j) \right) \quad (6)$$

where the weight  $\omega(x_i, x_j)$  indicates the structural similarity based on the patches  $P(x_i)$  and  $P(x_j)$  centred at  $x_i$  and  $x_j$ . Please note that it is impossible to use the Euclidean distance as a weight, as in the NL-means method, since the classified image is different from the grey-level value in an intensity image, i.e., it is not meaningful to assign a weight to a classification label. In this context, therefore, two weighting schemes are proposed for describing the structural similarity between pixels in a classified image.

(1) Consistency criterion (SSV-C). The consistency criterion is defined as

$$\omega(x_i, x_j) = \sum_v \text{eq}(P_v(x_i), P_v(x_j)) \quad (7)$$

where  $P_v(x_i)$  and  $P_v(x_j)$  are the pairwise pixels at the same location  $v$  within the patches centred at  $x_i$  and  $x_j$ . In this method, the voting weight is actually the number of pairwise pixels that have the same label.

(2) Histogram similarity criterion (SSV-H). This strategy is carried out in two steps. First, the local label histogram  $H(x_i)$ , which represents the frequency of each label occurring in patch  $P(x_i)$  centred at pixel  $x_i$ , is calculated. The weight is then defined as the similarity degree of the histograms. In this study, the intersection is used to measure the histogram similarity (Cha and Srihari 2002):

$$\omega(x_i, x_j) = \sum_{k=1}^m \min(H_k(x_i), H_k(x_j)) \quad (8)$$

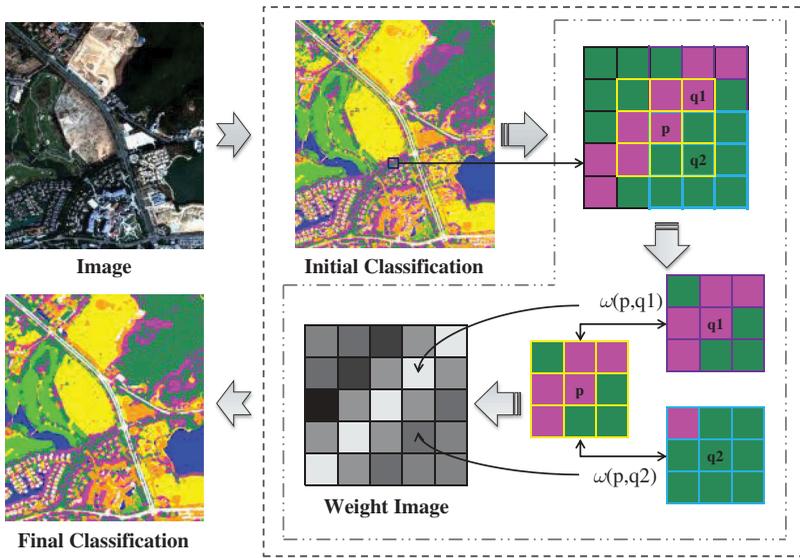
where  $H_k(x_i)$  represent the  $k$ th bin of the local label histogram for pixel  $x_i$ .

Figure 1 demonstrates the flowchart of the proposed algorithm. Here,  $q_1$  and  $q_2$  are pixels located in the window centred at pixel  $p$ . The pixels  $q_1$  and  $q_2$  have the same distance to the centre pixel  $p$ , but the neighbourhood of  $q_1$  is more similar to that of  $p$ , and therefore the weight  $\omega(p, q_1)$  is higher than  $\omega(p, q_2)$ . Correspondingly, the brightness (weight value) of  $q_1$  is greater than that of  $q_2$ .

### 3. Experiments

#### 3.1. Datasets and parameters

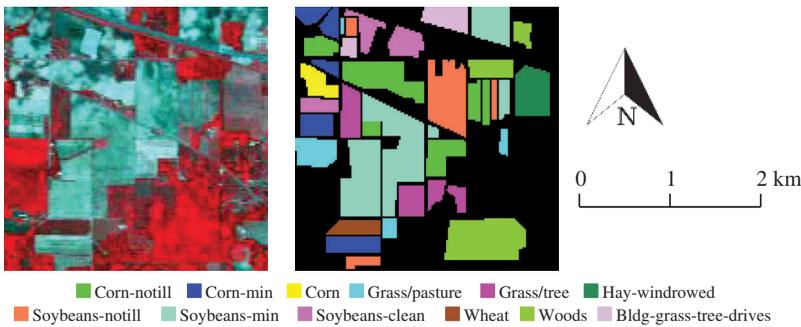
We investigated the effectiveness of the proposed approach with one hyperspectral and two high-resolution multispectral images. The hyperspectral image was acquired over



**Figure 1.** Flowchart of the proposed classification post-processing algorithm.

the Indian Pines test site in northwestern Indiana, USA, by the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS). The images and the reference data are shown in [Figure 2](#). The image is composed of  $145 \times 145$  pixels and 220 spectral bands (ranging from 1.4 to 2.5  $\mu\text{m}$ ), with a spatial resolution of 20 m. The image contains 10 classes, as reported in [Table 1](#). The first high-resolution image was acquired over Hainan Province, China, by the WorldView-2 satellite. The image is composed of  $600 \times 520$  pixels at a spatial resolution of 2 m, with eight spectral channels. The other high-resolution image was acquired over the city of Wuhan in Hubei Province, China, by the ZY-3 sensor, with four bands and  $651 \times 499$  pixels at a spatial resolution of 5.8 m. [Figure 3](#) presents the true colours and test references of the WorldView-2 and ZY-3 images. The numbers of test samples for these data sets are shown in [Table 2](#).

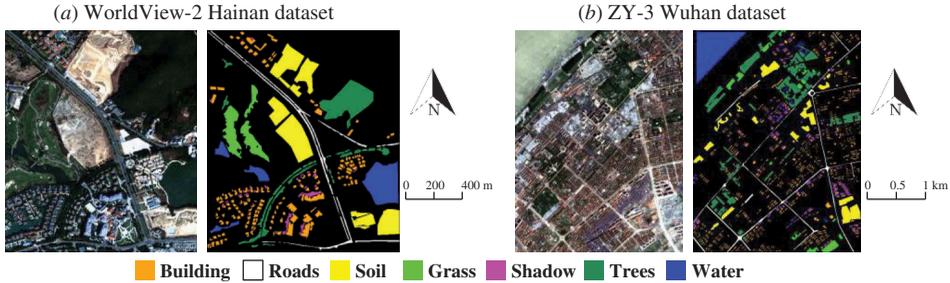
In this study, the raw classification results were obtained by support vector machine (SVM) with a radial basis function (RBF) kernel (Mountrakis, Im, and Ogole 2011). The penalty coefficient and bandwidth of RBF-SVM were set to the optimal values obtained



**Figure 2.** Hyperspectral image acquired by AVIRIS over the Indian Pines test site and the corresponding reference samples.

**Table 1.** Number of test samples for the hyperspectral AVIRIS data set.

Corn-notill	Corn-min	Corn	Grass/Pasture	Grass/Trees	Hay-windrowed
1434	834	234	497	747	489
Soybeans-notill	Soybeans-min	Soybeans-clean	Wheat	Woods	Bldg-Grass-Tree-Drives
968	2468	614	212	1294	380

**Figure 3.** High spatial resolution data sets. (a) The WorldView-2 Hainan image and the reference samples. (b) The ZY-3 Wuhan image and the reference samples.**Table 2.** Number of test samples for the WorldView-2 and ZY-3 data sets.

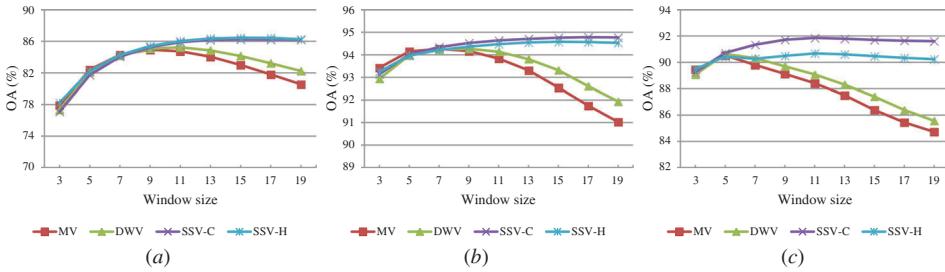
	Building	Roads	Soil	Grass	Shadow	Trees	Water
WorldView-2	10065	5932	8904	3321	3629	15169	11898
ZY-3	11578	5356	22189	7417	1427	14086	11209

by cross-validation. For the OBV approach, an adaptive mean-shift procedure (Huang and Zhang 2008) was employed for the segmentation. Fifty samples per class were randomly selected from the ground truth for training the classifier. To validate the robustness of the proposed methods, the experiments were repeated 10 times with randomly selected training samples.

### 3.2. Experimental results

Figure 4 shows the classification accuracies yielded by the filtering-based post-processing algorithms, where the horizontal and vertical axes indicate the size of the spatial window and the corresponding overall accuracy (OA). For all the data sets, the algorithms obtain similar results with a small window size, but the proposed method shows a significant advantage over the traditional ones in terms of classification accuracy when the window size increases. In other words, the proposed method is more robust to the window size than the traditional methods. This phenomenon can be attributed to the fact that the voting weights are determined by the structural similarities between the neighbourhoods of pixels rather than the spatial distance, which exploits the contextual information in the image more effectively.

The quantitative results for the test data sets are reported in Table 3. Here, the classification accuracies of the filtering-based approaches were obtained under a window size of 13. We can observe that the traditional methods can yield satisfactory results and improve on the raw classification accuracy by incorporating the neighbourhood



**Figure 4.** The overall accuracies (OA) obtained with different sizes of spatial window for the different algorithms tested. (a) the AVIRIS Indian Pines image, (b) the WorldView-2 Hainan image and (c) the ZY-3 Wuhan image.

**Table 3.** Classification accuracies (%) obtained by different algorithms for the AVIRIS Indian Pines image, the WorldView-2 Hainan image and the ZY-3 Wuhan image.

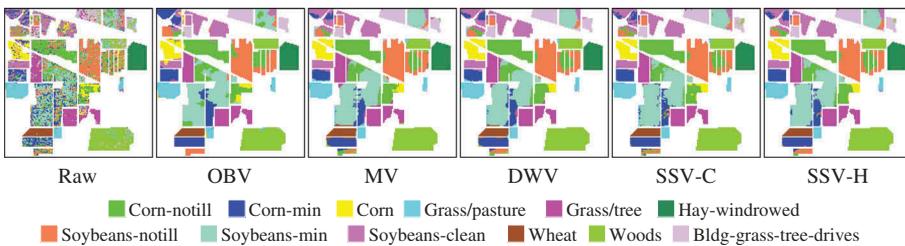
	Raw classification	Classification post-processing algorithms				
		OBV	MV	DWV	SSV-C	SSV-H
AVIRIS Indian Pines	68.28	83.33	84.04	84.85	86.17	86.37
WorldView-2 Hainan	91.45	93.43	93.31	93.81	94.71	94.55
ZY-3 Wuhan	86.87	89.46	87.51	88.31	91.79	90.62

information. It can also be concluded that the performance of the proposed algorithm is superior to the traditional approaches in terms of the classification accuracy.

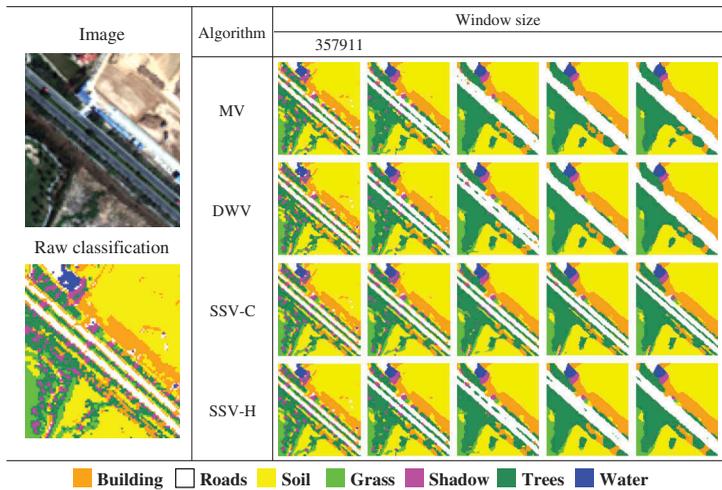
### 3.3. Visual interpretation

For visual interpretation, the classification maps of the AVIRIS image obtained by the different approaches are shown in Figure 5. Here, it can be clearly seen that the post-processing approaches effectively alleviate the salt-and-pepper misclassification errors and yield much smoother results.

Moreover, Figure 6 presents the classification maps of a sub-image from the WorldView-2 data set. Here, it can be seen that the filtering-based post-processing methods show similar results with a small window, while their results are quite different when the window size increases. This phenomenon is also supported by the quantitative evaluation. For the narrow and elongated green belt located between the roads, it can be seen that it is filtered out by the traditional post-processing approaches (MV and



**Figure 5.** Classification results of the different algorithms for the AVIRIS Indian Pines image.



**Figure 6.** Examples of the filtering-based classification post-processing methods under different window sizes for a sub-image from the WorldView-2 data set.

DWV); however, it is persevered by the proposed SSV-C method by effectively considering the contextual information for the filtering.

#### 4. Conclusion

In this article, a novel structural similarity-based label-smoothing algorithm has been proposed for the post-processing of remote-sensing classification. Based on the structural similarity between pixels, the voting weights for each neighbouring pixel are calculated adaptively. Specifically, a consistency criterion and a histogram similarity criterion are utilized to measure the structural similarity between pixels in the classified images. The experiments conducted on three data sets confirm that the proposed approach improves the initial classification accuracy and is able to give promising classification results. Compared to the traditional classification post-processing methods, the proposed algorithm shows its superiority in preserving the detail and structural information while smoothing the labelled image.

#### Disclosure statement

No potential conflict of interest was reported by the authors.

#### Funding

This work was supported by the National Natural Science Foundation of China under Grant 91338111, by China National Science Fund for Excellent Young Scholars under Grant 41522110, and by the Foundation for the Author of National Excellent Doctoral Dissertation of PR China (FANEDD) under Grant 201348.

## References

- Buades, A., B. Coll, and J.-M. Morel. 2005. "A Non-Local Algorithm for Image Denoising." Paper presented at the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005. CVPR 2005, San Diego, CA, June 20–25.
- Cha, S.-H., and S. N. Srihari. 2002. "On Measuring the Distance between Histograms." *Pattern Recognition* 35: 1355–1370. doi:[10.1016/S0031-3203\(01\)00118-2](https://doi.org/10.1016/S0031-3203(01)00118-2).
- Ghamisi, P., M. D. Mura, and J. A. Benediktsson. 2015. "A Survey on Spectral–Spatial Classification Techniques Based on Attribute Profiles." *IEEE Transactions on Geoscience and Remote Sensing* 53: 2335–2353. doi:[10.1109/tgrs.2014.2358934](https://doi.org/10.1109/tgrs.2014.2358934).
- Huang, X., Q. Lu, L. Zhang, and A. Plaza. 2014. "New Postprocessing Methods for Remote Sensing Image Classification: A Systematic Study." *IEEE Transactions on Geoscience and Remote Sensing* 52: 7140–7159. doi:[10.1109/tgrs.2014.2308192](https://doi.org/10.1109/tgrs.2014.2308192).
- Huang, X., and L. Zhang. 2008. "An Adaptive Mean-Shift Analysis Approach for Object Extraction and Classification From Urban Hyperspectral Imagery." *IEEE Transactions on Geoscience and Remote Sensing* 46: 4173–4185. doi:[10.1109/tgrs.2008.2002577](https://doi.org/10.1109/tgrs.2008.2002577).
- Liu, D., and F. Xia. 2010. "Assessing Object-based Classification: Advantages and Limitations." *Remote Sensing Letters* 1: 187–194. doi:[10.1080/01431161003743173](https://doi.org/10.1080/01431161003743173).
- Lu, Q., X. Huang, and L. Zhang. 2014. "A Novel Clustering-Based Feature Representation for the Classification of Hyperspectral Imagery." *Remote Sensing* 6: 5732–5753. doi:[10.3390/rs6065732](https://doi.org/10.3390/rs6065732).
- Ma, J., J. Zhao, and A. L. Yuille. 2016. "Non-Rigid Point Set Registration by Preserving Global and Local Structures." *IEEE Transactions on Image Processing* 25: 53–64. doi:[10.1109/tip.2015.2467217](https://doi.org/10.1109/tip.2015.2467217).
- Mountrakis, G., J. Im, and C. Ogole. 2011. "Support Vector Machines in Remote Sensing: A Review." *ISPRS Journal of Photogrammetry and Remote Sensing* 66: 247–259. doi:[10.1016/j.isprsjprs.2010.11.001](https://doi.org/10.1016/j.isprsjprs.2010.11.001).
- Myint, S. W., P. Gober, A. Brazel, S. Grossman-Clarke, and Q. Weng. 2011. "Per-Pixel Vs. Object-Based Classification of Urban Land Cover Extraction Using High Spatial Resolution Imagery." *Remote Sensing of Environment* 115: 1145–1161. doi:[10.1016/j.rse.2010.12.017](https://doi.org/10.1016/j.rse.2010.12.017).
- Schindler, K. 2012. "An Overview and Comparison of Smooth Labeling Methods for Land-Cover Classification." *IEEE Transactions on Geoscience and Remote Sensing* 50: 4534–4545. doi:[10.1109/tgrs.2012.2192741](https://doi.org/10.1109/tgrs.2012.2192741).