ISPRS Journal of Photogrammetry and Remote Sensing 141 (2018) 137–147

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

journal homepage: www.elsevier.com/locate/isprsjprs

Review Article

A review of accuracy assessment for object-based image analysis: From per-pixel to per-polygon approaches

Su Ye *, Robert Gilmore Pontius Jr., Rahul Rakshit

Graduate School of Geography, Clark University, Worcester, MA 01610, USA

ARTICLE INFO

Article history:
Received 3 January 2018
Received in revised form 4 March 2018
Accepted 5 April 2018

Keywords:
Accuracy assessment
Object-based image analysis
OBIA
Remote sensing
Per-pixel
Per-polygon

ABSTRACT

Object-based image analysis (OBIA) has gained widespread popularity for creating maps from remotely sensed data. Researchers routinely claim that OBIA procedures outperform pixel-based procedures; however, it is not immediately obvious how to evaluate the degree to which an OBIA map compares to reference information in a manner that accounts for the fact that the OBIA map consists of objects that vary in size and shape. Our study reviews 209 journal articles concerning OBIA published between 2003 and 2017. We focus on the three stages of accuracy assessment: (1) sampling design, (2) response design and (3) accuracy analysis. First, we report the literature's overall characteristics concerning OBIA accuracy assessment. Simple random sampling was the most used method among probability sampling strategies, slightly more than stratified sampling. Office interpreted remotely sensed data was the dominant reference source. The literature reported accuracies ranging from 42% to 96%, with an average of 85%. A third of the articles failed to give sufficient information concerning accuracy methodology such as sampling scheme and sample size. We found few studies that focused specifically on the accuracy of the segmentation. Second, we identify a recent increase of OBIA articles in using per-polygon approaches compared to per-pixel approaches for accuracy assessment. We clarify the impacts of the per-pixel versus the per-polygon approaches respectively on sampling, response design and accuracy analysis. Our review defines the technical and methodological needs in the current per-polygon approaches, such as polygon-based sampling, analysis of mixed polygons, matching of mapped with reference polygons and assessment of segmentation accuracy. Our review summarizes and discusses the current issues in object-based accuracy assessment to provide guidance for improved accuracy assessments for OBIA.

1. Introduction

1.1. Accuracy assessment for object-based image analysis

Object-based image analysis (OBIA) is quickly gaining popularity as a mapping technique among remote sensing researchers (Carleer et al., 2005; Mallinis et al., 2008; Peña-Barragán et al., 2011; Pu and Landry, 2012; Hussain et al., 2013; Blaschke et al., 2014; Robson et al., 2015). Previous non-OBIA methods characterized land cover on a pixel-by-pixel basis. In contrast, OBIA first divides images into homogeneous clusters of pixels called segments, also known as image objects (IOs) (Baatz et al., 2008; Addink et al., 2012; Pu and Landry, 2012; Sherba et al., 2014). Then, these IOs are applied as the spatial unit, instead of pixels, for image analysis such as image classification.

A wide body of remote sensing literature has claimed that OBIA techniques outperform traditional pixel-based methods (Cleve et al., 2008; Jobin et al., 2008; Walsh et al., 2008; Myint et al., 2011). The claimed improvement in OBIA classification accuracy has three arguments: (1) OBIA minimizes within-class spectral variability by assigning all pixels in the object to an identical land category (Peña-Barragán et al., 2011; Robson et al., 2015); (2) OBIA makes better use of spatial information implicit within remotely sensed images such as size, shape and texture of objects (Blaschke et al., 2014); (3) OBIA facilitates integration of contextual and semantic relationships among geographic objects (Platt and Rapoza, 2008; Blaschke, 2010; Tiede et al., 2010; Blaschke et al., 2014).

Several factors affect the quality of OBIA maps and can hinder the full potential use of remotely-sensed products. First, the step of segmentation for generating spatial objects for classification is an ill-posed problem (Hay and Castilla, 2008). Most segmentation algorithms require a trial-and-error process, which is often

* Corresponding author.
E-mail address: remotesensingsuy@gmail.com (S. Ye).

https://doi.org/10.1016/j.isprsjprs.2018.04.002
0924-2716/Published by Elsevier B.V. on behalf of International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS).

subjective and inconsistent (Arvor et al., 2013). Second, feature selection, which is the process of selecting relevant features to reduce the dimensionality for classification (Zhou and Wang, 2015), becomes tricky due to numerous object-specific features such as size, shape and perimeter of a polygon. These various features contribute to complexity of an OBIA task (Ma et al., 2017). Third, an OBIA classifier can use expert knowledge and semantics of image objects (Arvor et al., 2013; Du et al., 2015; Gu et al., 2017; Zhang et al., 2017), but the difficulties of establishing linkages between image objects and real-world geographical features may produce uncertainties and errors in the final maps (Blaschke et al., 2008). It is essential to have a meaningful and reliable way to evaluate OBIA results due to the abovementioned challenges (Stow et al., 2008; Blaschke, 2010). Errors and uncertainties propagate through analyses that link the map to other GIS-based applications (Arbia et al., 1998).

1.2. Per-pixel versus per-polygon approaches to accuracy assessment

We categorize approaches as either per-pixel or per-polygon based on the assessment unit, i.e., on whether pixels or polygons serve as the definition of agreement and disagreement with reference data (Stehman and Wickham, 2011). Per-pixel approaches aim to characterize accuracy in terms of area correctly classified and area misclassified; per-polygon approaches focus on the polygons or objects mapped as countable objects and quantify the number of objects classified correctly (Stehman and Wickham, 2011). The assessment unit is not necessarily identical to the sampling unit for a procedure of accuracy assessment. The sampled units can be aggregated or disaggregated for the final accuracy calculation. For example, a user might use polygons as the unit during sampling, but then analyze the individual pixels during the accuracy assessment, in which case we would categorize the accuracy approach as pixel-based. Later sections of our review discuss the effects of assessment units on sampling.

If one uses the per-pixel approach, then one must convert the OBIA polygon-based map into a raster map, and then sample the raster map. The rationale for using the per-pixel approach is that it provides area-based accuracy, which is of great interest to map users; furthermore it is relatively easy to use pixels for sampling designs (Stehman and Wickham, 2011). However, it is often difficult to align one pixel on a map to the exact same position in the reference data, even with Global Positioning Systems. Thus, Congalton and Green (2008) critiqued the use of a pixel as an assessment unit because a pixel is sensitive to positional errors. Furthermore, a pixel does not have a meaningful relationship with earth's features (Whiteside et al., 2014).

On the other hand, the per-polygon approach considers the thematic and geometrical properties of map units, i.e., polygons for OBIA. Therefore, Congalton and Green (2008) suggested "if the map to be assessed is a polygon map, then the accuracy assessment sample units should also be polygons". Some researchers claimed that fewer sampling units are needed for polygon-based sampling than for pixel-based sampling, thereby reducing costs of sampling (Radoux et al., 2011). However, the benefits of the per-polygon approach are still debatable. For example, Stehman and Wickham (2011) asserted that polygons represent real features of the landscape only if the polygons are defined in the reference information, regardless of the units in the OBIA map. Scientists must be aware of the impacts of the choice of per-pixel versus per-polygon approaches on accuracy assessments in order to improve practices and to interpret accuracy results.

Previous reviews surveyed the characteristics of OBIA for remote sensing applications, such as change detection (Hussain et al., 2013), supervised land-cover classification (Beckschäfer, 2017), wetland-specific research (Dronova, 2015), and ontological knowledge representation (Arvor et al., 2013). Our review is different from previous reviews because we examine accuracy assessment in OBIA. We focus on two objectives: (1) to identify the general characteristics of accuracy assessments for OBIA in terms of sampling design, response design, and accuracy analysis; (2) to review various implementations of per-pixel and per-polygon approaches in order to identify related controversial methodological issues.

2. Methods and data

2.1. Selection of articles

We examined peer-reviewed journal articles published from January 2003 to June 2017. The literature we found concerning OBIA published before 2003 were mostly conference proceedings, which we did not include in our review. We visited the websites of influential remote sensing journals and searched for articles using three keywords: OBIA, GEOBIA and object-based. We focused on OBIA articles concerning land cover/use classification, because it is critical for those articles that the geometric and thematic quality of their classified maps are evaluated in an appropriate manner. Our survey does not include articles that focused on object-oriented GIS analysis (Bian, 2007), or on the process of generating objects, i.e. segmentation. We found 209 articles, which serve as the basis for meta-analysis. Eighty-one percent of the selected articles (169) came from seven top remote sensing journals: Remote Sensing of Environment (39), Remote Sensing (38), ISPRS Journal of Remote Sensing (28), International Journal of Remote Sensing (25), Photogrammetric Engineering & Remote Sensing (15), International Journal of Applied Earth Observation and Geoinformation (16), and Remote Sensing Letters (8). The remaining articles (40) are from other relevant journals such as Canadian Journal of Remote Sensing, IEEE Transactions on Geoscience and Remote Sensing, Landscape and Urban Planning, Sensors, Applied Geography, etc. Our manuscript's supplemental materials give a complete list of the articles.

2.2. Selection of survey attributes

Table 1 describes 15 attributes that we recorded for the 209 articles. Attributes 1–8 are the primary attributes that we sur-

Table 1
Fifteen attributes for the reviewed articles.

Group	Attribute
Classification Design	1. Mapping Theme: impervious surfaces, vegetation, wetland, no specific theme or other 2. Type of remote sensing images: fine, moderate, and coarse grain.
Sampling Design	3. Sampling scheme: random, stratified random, systematic random, unspecified, or census. 4. Minimum per-category sample size
Response Design	5. Reference Data Source: field data, other GIS maps, office interpretation of remotely sensed images
Accuracy Analysis	6. Overall accuracy of the classification in percentage 7. Whether used confusion matrix: Yes or No 8. Whether performed segmentation accuracy assessment
Others	9. Year of publication 10. Name of Journal 11. Assessment approach: per-pixel or per-polygon 12. Method to match reference polygons with mapped polygons for per-polygon approach 13. Method to deal with mixed polygons 14. Segmentation assessment methods 15. Other accuracy assessment techniques used