

Article

Ontology-Guided Image Interpretation for GEOBIA of High Spatial Resolution Remote Sense Imagery: A Coastal Area Case Study

Helingjie Huang ¹, Jianyu Chen ^{1,*}, Zhu Li ¹, Fang Gong ¹ and Ninghua Chen ²

¹ State Key Laboratory of Satellite Ocean Environment Dynamics, Second Institute of Oceanography, State Oceanic Administration, Hangzhou 310012, China; hhlj@live.cn (H.H.); gzstrxx@163.com (Z.L.); gongfang@sio.org.cn (F.G.)

² Earth Science collage, Zhejiang University, Hangzhou 310012, China; geo316@zju.edu.cn

* Correspondence: chenjianyu@sio.org.cn

Academic Editor: Wolfgang Kainz

Received: 24 January 2017; Accepted: 29 March 2017; Published: 31 March 2017

Abstract: Image interpretation is a major topic in the remote sensing community. With the increasing acquisition of high spatial resolution (HSR) remotely sensed images, incorporating geographic object-based image analysis (GEOBIA) is becoming an important sub-discipline for improving remote sensing applications. The idea of integrating the human ability to understand images inspires research related to introducing expert knowledge into image object-based interpretation. The relevant work involved three parts: (1) identification and formalization of domain knowledge; (2) image segmentation and feature extraction; and (3) matching image objects with geographic concepts. This paper presents a novel way that combines multi-scaled segmented image objects with geographic concepts to express context in an ontology-guided image interpretation. Spectral features and geometric features of a single object are extracted after segmentation and topological relationships are also used in the interpretation. Web ontology language–query language (OWL-QL) formalize domain knowledge. Then the interpretation matching procedure is implemented by the OWL-QL query-answering. Compared with a supervised classification, which does not consider context, the proposed method validates two HSR images of coastal areas in China. Both the number of interpreted classes increased (19 classes over 10 classes in Case 1 and 12 classes over seven in Case 2), and the overall accuracy improved (0.77 over 0.55 in Case 1 and 0.86 over 0.65 in Case 2). The additional context of the image objects improved accuracy during image classification. The proposed approach shows the pivotal role of ontology for knowledge-guided interpretation.

Keywords: GEOBIA; ontology; multi-scale image objects; context

1. Introduction

Geographic object-based image analysis (GEOBIA) is recognized as an evolving paradigm in the remote sensing image-processing domain [1]. It consists of image segmentation and subsequent analysis of the image object. In past decades, many algorithms have been proposed for obtaining outstanding segmentation results while endeavoring to achieve optimal segment and potential real image object interpretation [2,3]. The GEOBIA paradigm continues to show its efficacy in remote sensing image analysis by providing tools that emulate human perception and combine an analyst's experience with meaningful image objects [4]. Few works have focused on object-based image interpretation. Obtaining subsequent information has generally relied on fuzzy- and/or role-based classification [5], which is also the main approach applied in the commercial software eCognition [6]. With increasing acquisition of large volumes of high spatial resolution (HSR) remote sensing images,

content-based modeling for image scene recognition is becoming more important and relevant. Knowledge representation techniques play a pivotal role in the future evolution of remote sensing [7].

Image interpretation can be described as the semantic extraction of an image [8]. It consists of obtaining useful spatial and thematic information on image objects using human knowledge and experience [9,10]. Ontology is a popular knowledge representation technology in information science. Ontology is a formal, explicit specification of a shared conceptualization [11], which formally names and defines the classes, properties, and interrelationships of entities in a particular domain. It requires constructing natural language semantics in a formalized logical expression that can be processed by a computer. Recently, the role of ontology in the interpretation of remote sensing has been highlighted; Andrés et al. [12] used spectral rules formalized in an ontology to identify the Brazilian Amazon area. In addition to spectral features and geometric features, Forestier et al. [13] added the descriptions of neighbor objects in order to interpret a coastal image. Luo et al. [14] used texture to classify land cover. Meanwhile, objects in urban settings have prominent geometric features, so these areas have received additional attention [8,15,16]. The difference between low-level features extracted from images and high-level geographic meaning from human cognition, known as the semantic gap, is the core problem in knowledge-guided interpretation. From the knowledge side, it is the expert's work to explore the features of a geographic concept. The other side is extraction and expression of image features. Linking more image features with expert knowledge will improve interpretation. The features of single image objects (e.g., spectral features, geometric features) have been used in previous studies. For improvement, additional features among image objects must be taken into consideration, such as the spatial pattern and context.

Here, our objective is to develop an ontology-based image object interpretation. This paper presents a novel method that combines multi-scaled segmented image objects with a geographic concept definition to express the context for better interpretation. The proposed approach is introduced in Sections 2.1–2.5. First, image segmentation work is performed using eCognition software, and the spectral features and geometric features of each segment are extracted. The hierarchical structure obtained from multi-scaled segmentation serves as the indirect expression of context. Besides spectral features, geometric features, and topological relationships are used. Ontology stores the geographic terms and their definitions in web ontology language–query language (OWL-QL). For efficiency, the match between concepts and objects is implemented by OWL-QL query-answering. Case studies utilizing this technique are illustrated in Section 3 along with results, assessment and discussion. Conclusions follow in Section 4.

2. Methodology

2.1. Ontology-Guided Image Interpretation for Image Object

A new ontology-guided approach for geographic object-based image interpretation is proposed for HSR images (Figure 1). The approach consists of three parts. The first part is image segmentation and features extraction, generating multi-scaled image objects, which are then evaluated by supervised image segmentation assessment showing the differences between segmented image objects and geographic objects of interest. The spectral and geometric features are calculated for each obtained image object. Instead of calculating in advance, the topological relationships and context are visited. The interpretation matching procedure queries a geographic concept to obtain corresponding image objects supported by the OWL-QL query-answering procedure.

Second, the knowledge of relevant geographic concepts is identified and then formalized in an ontology. OWL-QL, a sublanguage of OWL for conducting a query-answering dialogue among ontologies [17], builds upon this ontology. By separating the conceptual knowledge from the factual knowledge, it applies to an ontology that has a large amount of factual knowledge. Spectral features, geometric features, and topological relationships as object property terms directly appear within the

ontology. With these features, combining the multi-scaled image objects defines geographic concepts. It is explained in detail in Section 2.4.2.

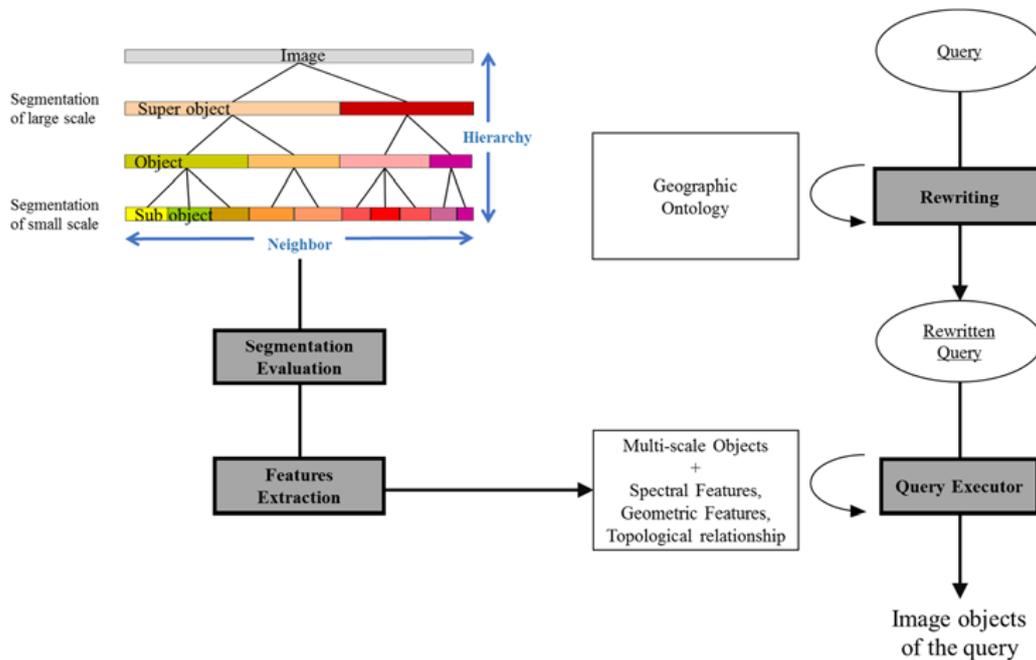


Figure 1. Workflow of ontology-guided interpretation for image objects.

The last part of this approach is the interpretation matching procedure, also called a classifier. Benefiting from OWL-QL, the factual knowledge can be separated from the conceptual knowledge for efficiency. The conjunctive query communicates between the two types of knowledge. The interpretation starts with a query of a geographic concept to obtain the type of image objects. Then, based on the ontology, the query is rewritten to a set of queries. They are derived from two parts: the definition of the queried geographic concept and the reasoned related knowledge. For instance, there is a query of the class of *Inward Flowing River* that is ontologically defined as a kind of river that does not flow into the sea. By the rewriting, the query can be extended to a query of *River* and a query of *Not Flowing Into The Sea*. Then the extension continues, adding a query of the *Sea* and a query of what objects do *Not Flowing Into It*, until the related concepts are exhaustively searched. In the rewriting, the reasoner works to take the implicit knowledge into consideration. Thus, a conceptual query through the geographic ontology gains knowledge to a set of queries. The interpretation ends with the set of queries answered by factual knowledge storage (e.g., database), obtaining the corresponding image objects.

2.2. Multi-Scaled Segmentation and Evaluation

Image segmentation is a fundamental step in GEOBIA. In segmentation, an image is divided into sets of contiguous pixels as image objects that are spectrally homogeneous inside and heterogeneous outside [18]. The degree of homogeneity determines the scale of image objects. Image segmentation, as an ill-posed problem, requires input parameters to be tuned by an expert, usually following a trial-and-error process [19] to obtain the optimal segmentation. In a widely-used segmentation, it is common to select a set of parameters to fit the geographic objects of interest, which is usually slightly over-segmented in practice.

Image segmentation influences the quality of the subsequent interpretation, because a segmented image object is the basic unit of analysis. Supervised image segmentation assessment measures the differences between the reference image object and the segmented image object. According to the geographic objects of interest, the researcher delineates the reference objects. Cheng et al. [20] deriving

from the object-fate analysis method, divided the segmented objects that intersect a reference object into three types: good objects, expanding objects, and invading objects. The good object is completely within the reference object. The intersecting area of the expanding object and reference object accounts for more than 50% of the area of the expanding object, while the intersecting area of the invading object is less than 50%. Good and expanding objects merge as the matched objects of the reference object. In this study, the differences are measured from three perspectives: quantity, area, and position.

Two quantity evaluations introduced here are by Schöpfer and Lang [21]: offspring loyalty (OL) and interference (I).

$$OL = \frac{n_{\text{good}}}{n_{\text{good}} + n_{\text{exp}}} \quad (1)$$

$$I = \frac{n_i}{n_{\text{all}}} \quad (2)$$

where n_{good} , n_{exp} , n_i , and n_{all} represent the number of good objects, expanding objects, invading objects, and all intersecting segmented objects, respectively.

Area evaluations are shown in Table 1. The Area Fitness Index (AFI) was proposed by Lucieer and Stein [22], and the remainder from Cheng et al. [20].

The Position Discrepancy Index (PDI) describes the average distance between the reference object and its matched image objects [20]. Overall PDI is the average of the PDI.

$$PDI = \frac{1}{N + M} \left(\sum_{k=1}^N \sqrt{(X(k) - X_r)^2 + (Y(k) - Y_r)^2} + \sum_{l=1}^M \sqrt{(X(l) - X_r)^2 + (Y(l) - Y_r)^2} \right) \quad (3)$$

$$PDI_{\text{overall}} = \frac{1}{n} \sum_{i=1}^n PDI(i) \quad (4)$$

where N and M are the number of good objects and expanding objects, respectively (the two are both matched objects), $(X(k), Y(k))$ is the centroid of the k -th good object, $(X(l), Y(l))$ is the centroid of the l -th expanding object, and (X_r, Y_r) is the centroid of the reference object.

Table 1. Area evaluation in the supervised assessment of image segmentation.

Measurement	Definition	Description
Area Fitness Index (AFI)	$\frac{A_r - A_{\text{Largest Image Object}}}{A_r}$	When AFI > 0, over segmentation; When AFI < 0, under segmentation
Omission Error (OE)	$\frac{\sum_{j=1}^{n_i} \{A_i(j) \cap A_r\}}{A_r}$	Describes the over-segmentation. An OE closer to zero means less over-segmentation.
Commission Error (CE)	$\frac{\sum_{k=1}^{n_{\text{exp}}} \{A_e(k) - (A_e(k) \cap A_r)\}}{A_r}$	Describes the under-segmentation. A CE closer to zero means less under-segmentation.
OE _{overall}	$\frac{\sum_{i=1}^n (\text{OE}(i) \times A_r(i))}{\sum_{i=1}^n A_r(i)}$	The weighted average of OE.
CE _{overall}	$\frac{\sum_{i=1}^n (\text{CE}(i) \times A_r(i))}{\sum_{i=1}^n A_r(i)}$	The weighted average of CE.
Overall Area Discrepancy Index (ADI _{overall})	$\sqrt{\text{OE}_{\text{overall}}^2 + \text{CE}_{\text{overall}}^2}$	The overall of over- and under-segmentation. When ADI is zero, the segmentation is exactly the objects of interest.

Note: A_r is the area of the reference object, and $A_{\text{Largest Image Object}}$ is the area of the largest segmented object in the intersecting objects of one reference object. $A_i(j)$ is the area of the i -th invading object, $A_e(k)$ is the area of the k -th expanding object. In addition, n is the number of reference objects.

2.3. Feature Extraction

Features are the bridge between facts and concepts. Facts are generalized into concepts through features, by which concepts classify facts. The geographic concepts are defined using the features of image objects for filling the semantic gap. Assuming that the knowledge is credible, the ontology guided interpretation becomes more capable, as image features bridge more knowledge.

In this study, three types of features define the geographic concept: features of a single image object, features between two image objects, and context. We first consider the features of a single image object, such as spectral features, geometric features, and texture. They are usually regarded as attributes, including qualitative and quantitative values. Spectral features (Figure 2) are used to recognize substances (water, vegetation, soil or sealed ground) as inherited from pixel-based interpretation methods. Geometric features (Figure 3) can contribute to the interpretation of additional information. For example, water can be a pond, river or lake according to the size and shape [23]. Topological relationships are important spatial relationships between two image objects. Because image objects in a segmentation of the same scale are seamless and non-overlapping polygons, topological relationships are reduced to adjacency. Between the segmentations of different scales, topological relationships are adjacent, contained or within. Context is the key for further improvement of interpretation. The meaning of an object is not just from itself, but also from the surroundings. The topological relationships contribute to this context. As for context involving multiple image objects, we express it via the hierarchy within multi-scaled image objects.

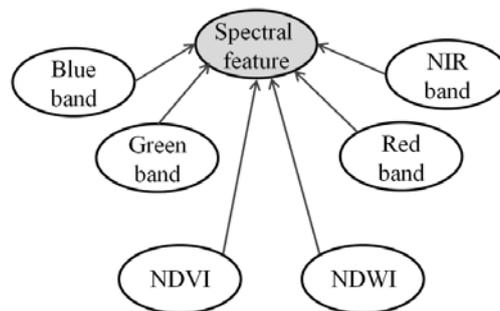


Figure 2. Spectral features.

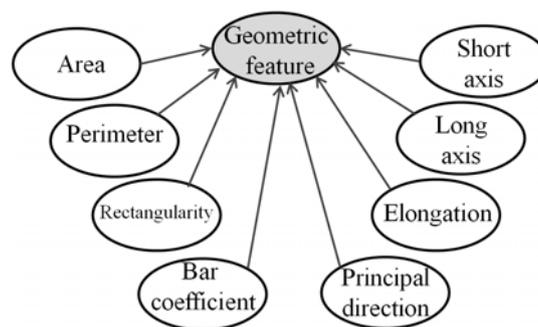


Figure 3. Geometric features.

In this study, the image objects are stored as polygons in ESRI Shapefile. The features of a single image object have a one-to-one relation to its image object. In this step, only the spectral features and geometric features are extracted and stored as attributes. The interpretation procedure queries topological relationships and context in real time.

2.4. Geographical Ontology for a Coastal Area

2.4.1. Ontology

Knowledge representation is aimed at constructing semantics in a computer readable manner and then processing it by programs. Ontology is a knowledge representation based on description logic. It is a formal naming and definition of the classes, properties, instances, and interrelationships of the entities for a particular domain. All descriptions can be connected to form a semantic graph that can be serialized using various formats, e.g., XML, turtle, N3 [19]. At present, OWL is a popular ontology language based on XML and standardized by W3C [24]. Ontology becomes powerful with the reasoner (e.g., HermiT, Pellet, FaCT++), which is used to check the logic consistency in the knowledge base and infer the implicit knowledge from the knowledge base. It highlights the capacity of ontology in knowledge management and knowledge discovery, which benefits a broad range of domains.

Ontology construction is complex and faced with two main challenges. One is concept identification and definition. It is the reason the technology that is named after the philosophical term ‘Ontology’. The other is the actual construction of the knowledge base [8]. In the geographic domain, it is an especially puzzling problem to identify concepts. Most geographic objects have qualitative descriptions instead of quantitative definitions. The same term may refer to something similar but show different features in different places. Additionally, the boundaries of natural geographic objects are often indeterminate. It is necessary to clarify the definition of terms in detail under a specific background. Therefore, knowledge formalized in geographic ontology is highly dependent on a specific application. For construction from natural language to formalized logical expression, the problem is general in knowledge engineering—formalizing the exact intended semantics and maintaining logic consistency. It forces the researcher to be the domain expert and the knowledge engineer at the same time. The test of logic consistency can rely on the reasoner, but semantics errors are often insidious.

2.4.2. Concept Definition Working with Multi-Scaled Image Objects

In ontology-guided interpretation of image objects, concept definition has to work with the image objects, even from the segmentation step. Humans create concepts to recognize reality by generalizing the features of an object and creating a definition. The image object acts as a medium in the interpretation: It is segmented consistently with the intended object. Through its features, the geographic concept interprets the image object as the corresponding geographic object. Both image segmentation and concept definition require an expert involved, so it is the expert’s work to make them cooperate. In this approach, we extend the cooperation to multi-scaled image objects for the expression of context.

The cooperation between the concept definition and the multi-scaled image objects is illustrated in Figure 4. For instance, Figure 4a is an image object from the level 2 segmentation in Figure 4b. It represents a typical seaside reclamation region that has several artificial ponds with banks by the sea. Previous methods usually segment it in one scale (as the yellow line shows corresponding the level 3 segmentation in Figure 4b) and then use spectral features to classify water and bare land, along with geometric features to further identify water as ponds (area and shape index) and bare land as roads (shape index). However, the arrangement and combination of image objects showing regional features are ignored. Context is key for better interpretation. Considering the surroundings, roads beside the artificial ponds can be recognized as pond banks. Inside the region, the artificial ponds and banks account for most of the area, and can be identified as a region of cultivation ponds. The ponds and banks within the region have further interpretation as cultivation ponds and banks respectively.

This approach proposes three ways of context supported by the topological relationships: (1) context from the surroundings (its neighbors); (2) context from the components (its sub-image objects); and (3) context from the region (its super-image objects). This method treats contiguous image objects that share spectral homogeneity at larger scale as associated objects. They become one object in a larger scale segmentation. Through the spatial relationship, contain and within, the sub-image

objects are combined to form context showing regional features. The super-object holds the scope of region. In other words, the super-object is interpreted by its own features and interior details are provided by the sub-objects, and in turn, add information to the sub-object's context. From bottom to top and returning to the bottom, a single image object is identified not only by itself (spectral and geometric features) but also by its context.

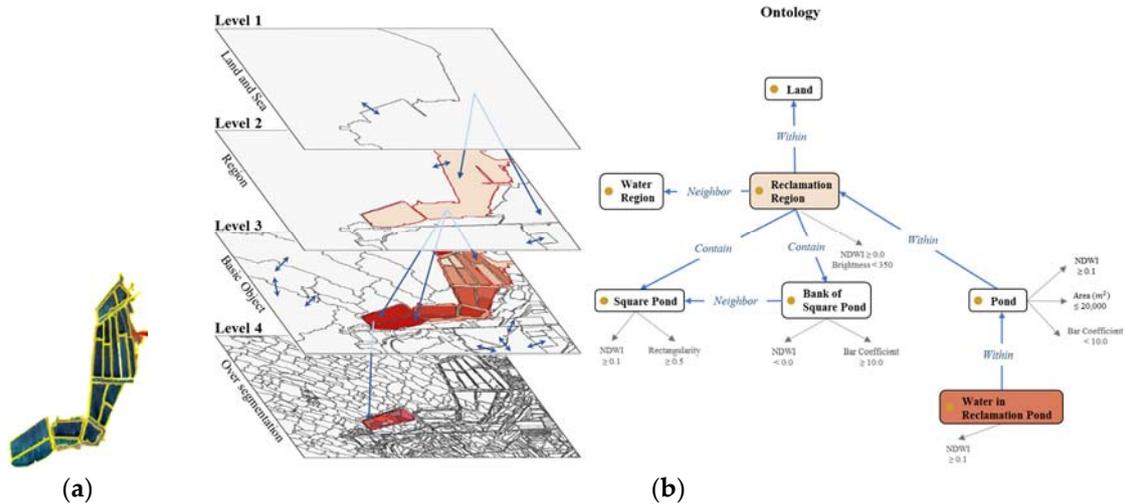


Figure 4. (a) Seaside reclamation region from the colored segment of Level 2; (b) The cooperation between the multi-scaled segmentation and geographic concepts.

2.5. OWL-QL Query and Answer

Some regard the reasoner as a special classifier, because of the ability to infer implicit knowledge. Reasoning is typically multi-exponential or even undecidable. Theoretically, a large number of conclusions are drawn from the ontology, but only some are needed in practice. In the case of remote sensing image interpretation from instances (image objects) to classes (geographic concepts), the large number of instances makes reasoning time consuming. OWL-QL, a sublanguage of OWL, supports query-answering within the ontology, and reasoning is performed in polynomial time [25] applied to the situation. Following the method of Krötzsch [25], there are three steps to complete the OWL-QL query-answering (Figure 5).

- The user specifies a query in the form of a conjunctive query, for instance, the query *WaterInReclamationPond(x)* to retrieve this kind of image objects.
- Using ontology that only contains concept descriptions, the query is rewritten into a set of queries still in the form of a conjunctive query, which means the query is extended by the ontology according to inference rules. This process is called rewriting-based reasoning.
- Rewritten queries are answered using the database or ontology that only stores the instances and its properties.

This framework takes concepts and instances apart to reduce the time complexity both when performing queries and reasoning. Another advantage is that the features are not necessarily extracted in advance, especially for features that have no one-to-one relation to its image object, such as topological relationships or context. They are visited when answering the related queries. OWL-QL is a lightweight language that sacrifices some expressiveness in comparison with OWL 2. Here, query-answering is based on REQUIEM (RESolution-based QUery rEWriting for Expressive Models) [26], a prototypical implementation of a query rewriting algorithm [27].

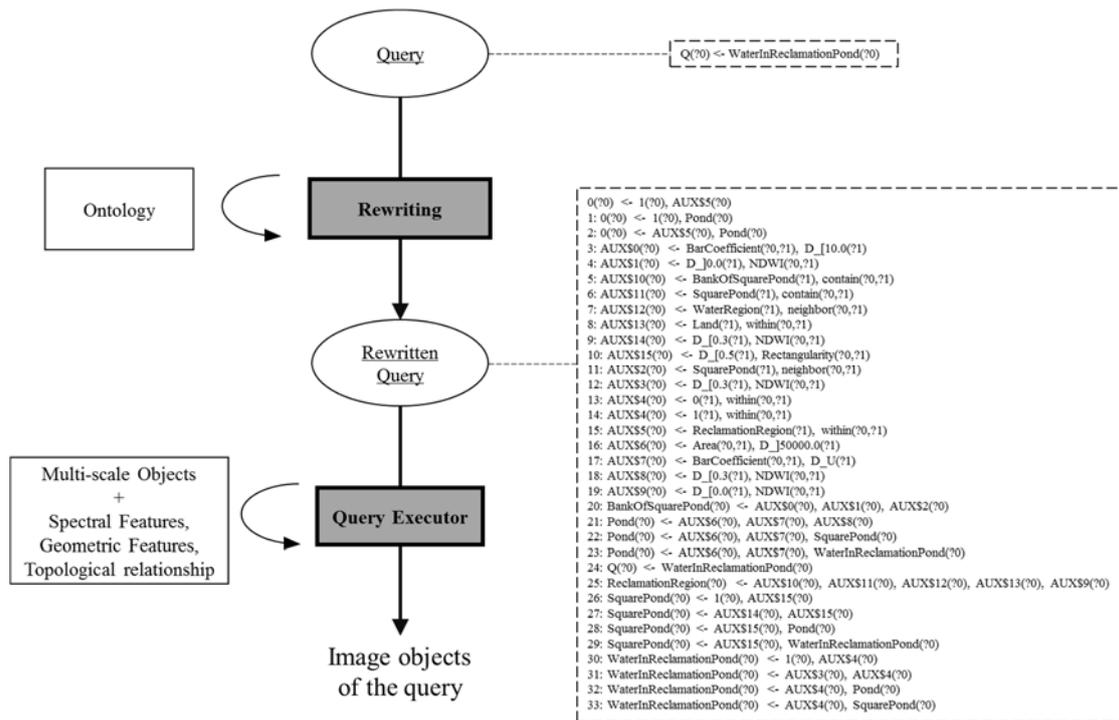


Figure 5. Conjunctive query rewriting and answer.

3. Case Study

3.1. Data

To illustrate using the proposed approach for advancing image object interpretation, experiments were carried out on two images of coastal districts in China (Figure 6). Four spectral bands of images were included: blue (Band 1), green (Band 2), red (Band 3), and near-infrared (Band 4).

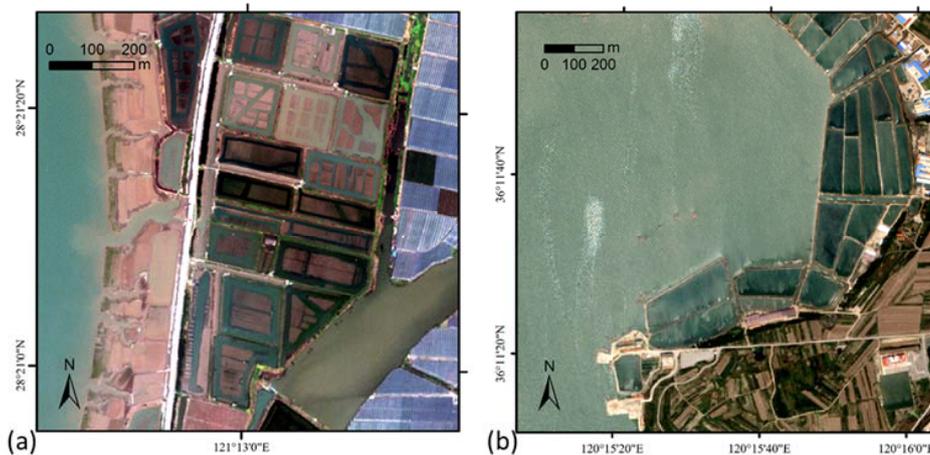


Figure 6. Experiment images with a composite of Bands 1, 2, and 3. (a) Image was acquired by Worldview2 on 31 March 2012; located in Yandangshan, China; 500 × 500 pixels with 2.4 m spatial resolution; (b) Image was acquired by Quickbird2 on 28 June 2008; located in Jiaozhouwan, China; 600 × 600 pixels with 2 m spatial resolution.

The two example image scenes, chosen for their prominent features and context, are composed mainly of water, bare land, mudflats, seaside reclamation, greenhouses, artificial structures, and fields. Seaside reclamation ponds are common in coastal areas, for various sea uses, such as aquaculture

or bay salt. The reclamation ponds have prominent features. From the photo of District 1 shown in Figure 7, reclamation ponds are arranged neatly adjacent to the sea or river. Due to their shallow depth, mudflats or bare land appear in the middle or at the edge of the ponds. The banks of the reclamation ponds are composed of bare land, occasionally covered with vegetation.

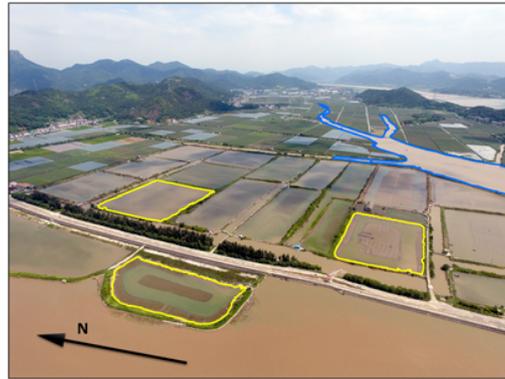


Figure 7. Photo of District 1 taken by an unmanned aerial vehicle.

3.2. Experiments and Discussion

The study started with image segmentation producing image objects on four scales. The commercial software eCognition conducted the multiresolution segmentation algorithm with four sets of parameters differing in scale. The choice of parameters depends on the four-scaled objects of interest: regions of sea and land (Figures 8a and 9a); large regions, such as water, mudflats, reclamation, and fields (Figures 8b and 9b); basic geographic objects of interest (Figures 8c and 9c); and over-segmented image objects for refined classification (Figures 8d and 9d). The image objects of the smallest scale can be regarded as the final classification units. Therefore, the image objects of Figures 8d and 9d were put into the supervised assessment of image segmentation.

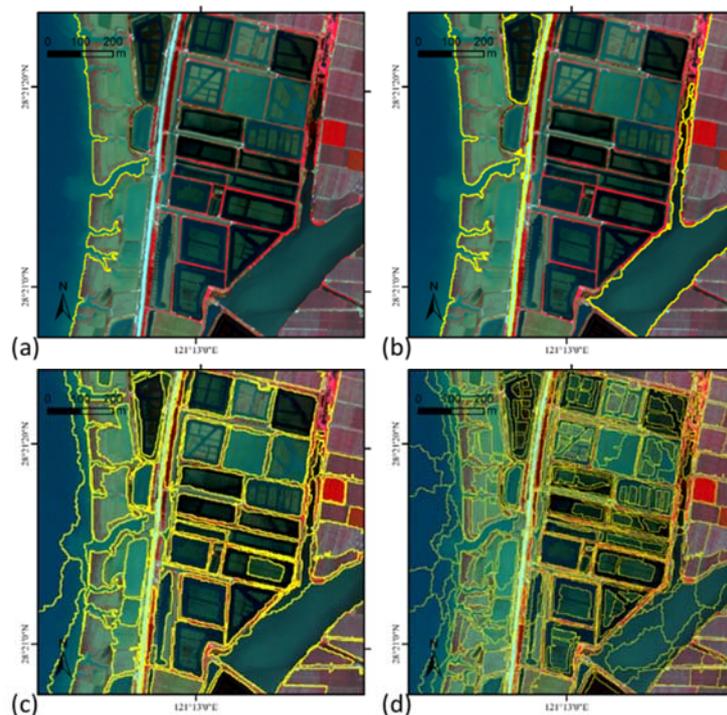


Figure 8. Segmentation results from Case 1 with segmentation parameters: shape as 0.1, compact as 0.9. The scale is 800 for (a); 500 for (b); 100 for (c); and 35 for (d).

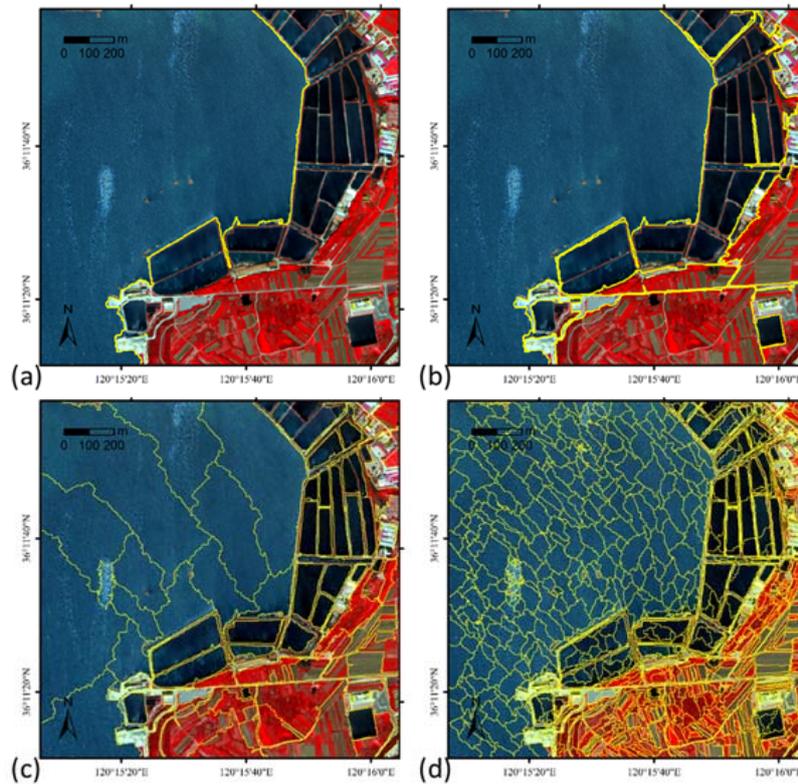


Figure 9. Segmentation results from Case 2 with segmentation parameters: shape as 0.6 and compact as 0.3. The scale is 700 for (a); 275 for (b); 100 for (c); and 20 for (d).

The supervised image segmentation assessment cannot represent the overall accuracy. Samples were selected to estimate segmentation quality by measuring the differences between the reference objects and the segmented objects. The bare land, greenhouse, water pond, vegetation, and mudflats are the objects of interest, so 10 reference objects were delineated for two images (Figures 10 and 11). The difference indices of quantity, area, and position (defined in Section 2.2) were calculated. There are some similarities in the two segmentation assessment results (Tables 2 and 3); all values of I are more than 0.5, and most values for OL are zero. The number of invading objects accounts for most of this result, and there are a few good objects. However, the low values of OE and CE show that the segmentation error is small and slightly over-segmented as a whole (all values of AFI are lower and higher than zero). This situation may be caused by the delineation of the reference objects. The object boundaries are presented by gradual mixing pixels so that precise delineation decisions are rare. The segmentation result is optimal when both $ADI_{overall}$ and $PDI_{overall}$ are at the minimum simultaneously [28]. This means the areas of commission and omission are small and the matched image objects are close to the reference image object. However, such a case rarely occurs. Area discrepancy information is more important than position discrepancy information for segmentation assessment [28]. The values of the two regions' $ADI_{overall}$ are 0.05 and 0.08. The segmentation result is valid and can be used for the subsequent process.

Table 3. Supervised assessment results of Case 2 (d) segmentation.

Reference Object	Description	n_{good}	$n_{\text{expanding}}$	n_{invading}	OL	I	AFI	OE	CE	PDI
0	Water	0	2	7	0	0.78	0.12	0.03	0.02	112.07
1	Water	0	6	8	0	0.57	0.71	0.02	0.03	62.56
2	Water	4	12	8	0.25	0.33	0.86	0.02	0.02	88.15
3	Bare land	0	3	8	0	0.73	0.47	0.08	0.00	39.97
4	Water	0	3	5	0	0.63	0.17	0.02	0.03	66.58
5	Bare land	0	1	19	0	0.95	0.19	0.95	0.04	176.97
6	Bare land	0	1	4	0	0.80	0.08	0.08	0.05	18.79
7	Structure	0	2	10	0	0.83	0.50	0.21	0.10	51.25
8	Bare land	0	2	7	0	0.78	0.47	0.04	0.11	27.39
9	Vegetation	0	3	5	0	0.63	0.53	0.03	0.11	35.73
Overall								0.07	0.03	67.94
Overall ADI	0.08									

The proposed approach interpreted the image objects in the two cases (Figures 12a, 13a and 14a). For accuracy, image objects were delineated and assigned types as reference interpretations (Figures 12c and 14c), and then they were formed as error matrices in pixels (Figures 13a and 15a). The overall accuracy and kappa were computed showing the errors including segmentation and interpretation. In order to show the capacity of the proposed approach, the supervised classification interpreted the image objects for comparison (Figures 12b and 14b). The eCognition software performed the supervised classification with the spectral features (mean values of Bands and MaxDifference) and geometric features (area, length, rectangular fit, roundness, and density). It also helped to show the role of context, because the supervised classification analyzes the features of single image objects, treating the image objects independently. Figures 13b and 15b shows the accuracy assessment.

In Case 1, the proposed approach identified 19 classes and the overall accuracy was 0.77 with a kappa of 0.73. Most of the area is well identified; the water-related concepts and greenhouse were clearly classified, because they have internal homogeneity and a distinct border. The mudflat as a whole shows the correct interpretation, but the mud land and bare land within the mudflat are muddled. In the reference interpretation, it is hard to distinguish mud land and bare land, especially in the transitional region. This situation also appears in the mud land and bare land that are within the reclamation ponds. Other errors are mainly from mixed pixels, which can be caused by segmentation, interpretation, and reference interpretation.

The supervised classification in Case 1 resulted in 10 classes with an overall accuracy of 0.55 and a kappa of 0.49. Two kinds of water are roughly separated—pond water and another large area of water. This is due to the spectral differences of water quality and area. The image objects near the shore, of what should be the large area of water, were wrongly identified as pond water for the same reason. Here, a more appropriate way to name the two types of water is *Clean Water* and *Shallow Water With A Mud Bed*. This case illustrates that spectral features that recognize the substances and geometric features can advance classification.

In Case 2, the proposed method obtained an interpretation that has 12 classes with an overall accuracy of 0.86 and a kappa of 0.79. The water region and reclamation were well interpreted; the result shows richer details than the reference interpretation in the regions of vegetation and bare land. As they are mixed, which led to errors in delineating image objects in the reference interpretation, several main roads were identified but misclassified. The sealed or unsealed roads were wrongly interpreted because flawed spectral rules cannot distinguish between them. Fragments of the roads were wrongly interpreted as bare land, an error similarly caused by mixed pixels.

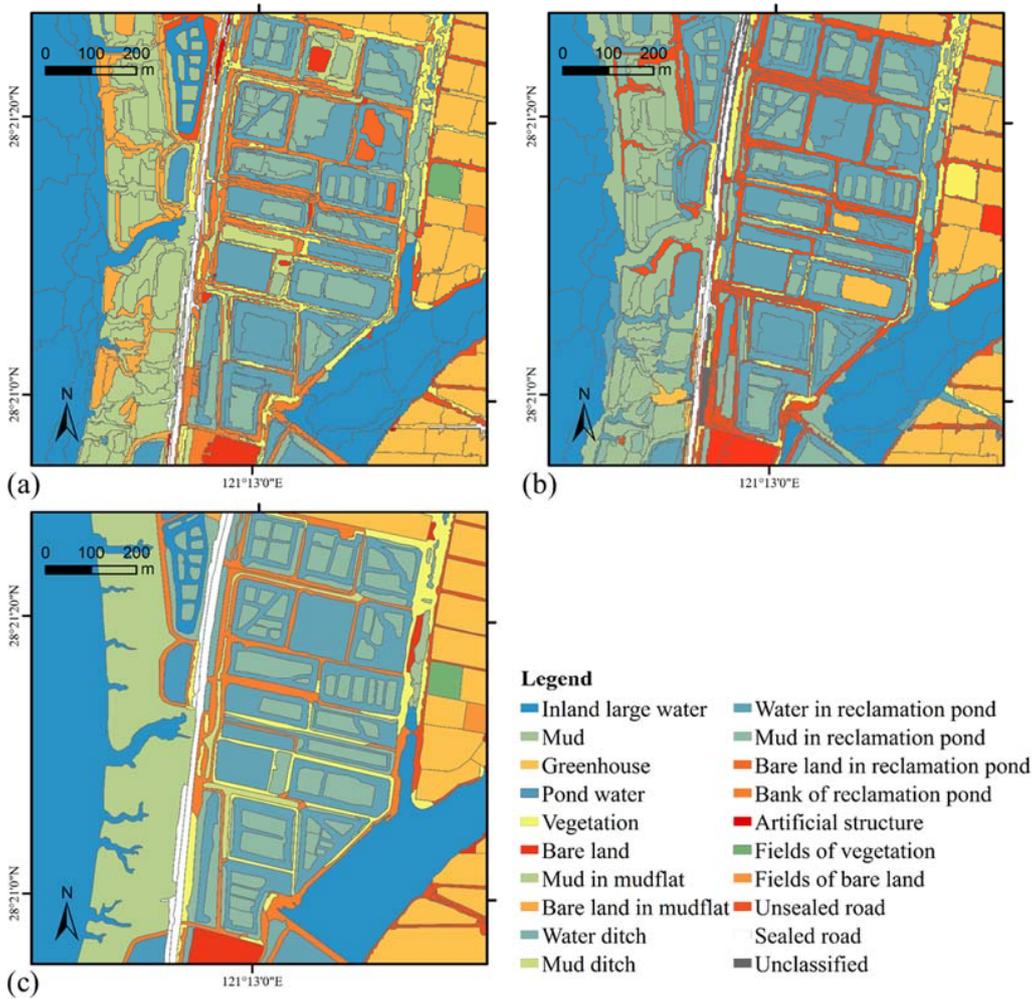


Figure 12. Case 1: (a) interpretation using the proposed method; (b) interpretation using the supervised classification; (c) reference interpretation.

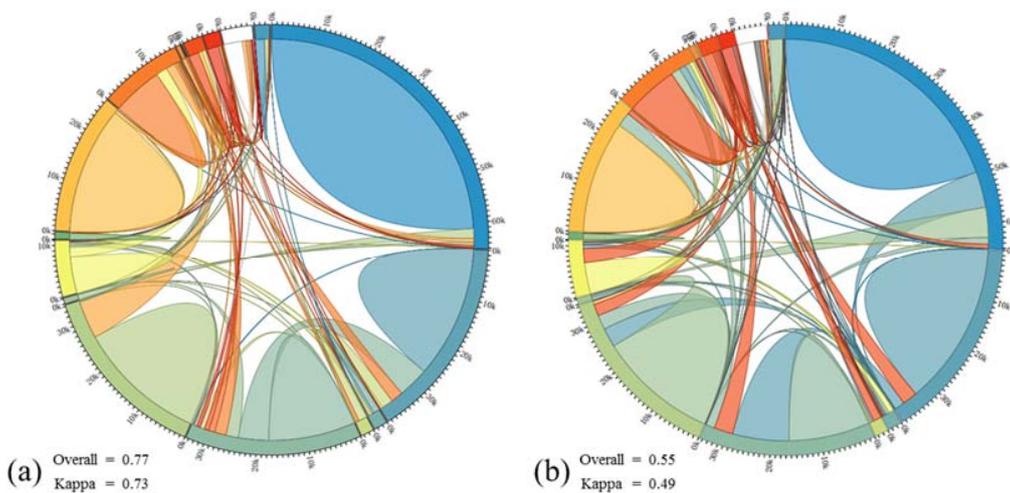


Figure 13. Chord diagram of the error matrix in Case 1: (a) between the reference and interpretation of the proposed method; (b) between the reference and interpretation of the supervised classification. The ribbons on the circle represent the classes of reference interpretation whose length is the number of pixels. The arches indicate the correct classification, and chords indicate incorrect classification (the legend in the diagram is the same as in Figure 12).

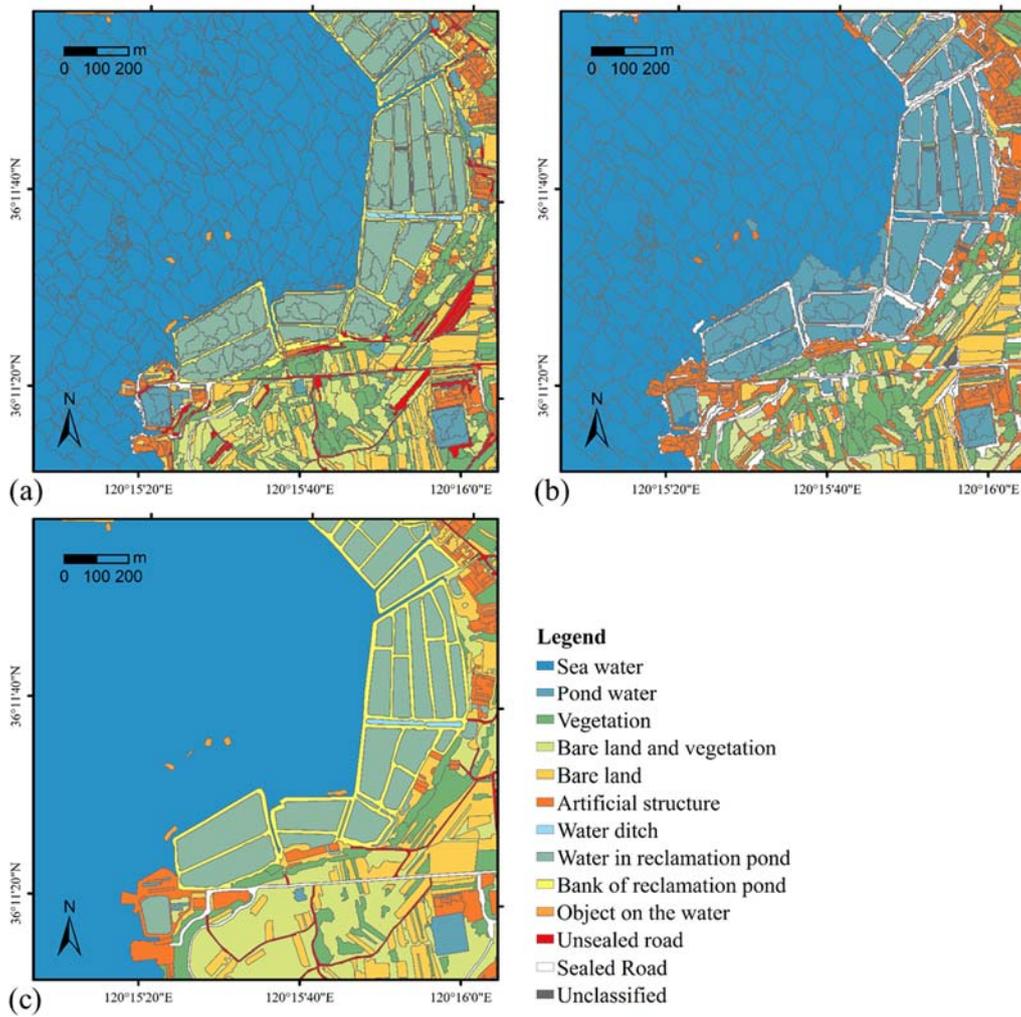


Figure 14. Case 2: (a) interpretation using proposed method; (b) interpretation using supervised classification; (c) reference interpretation.

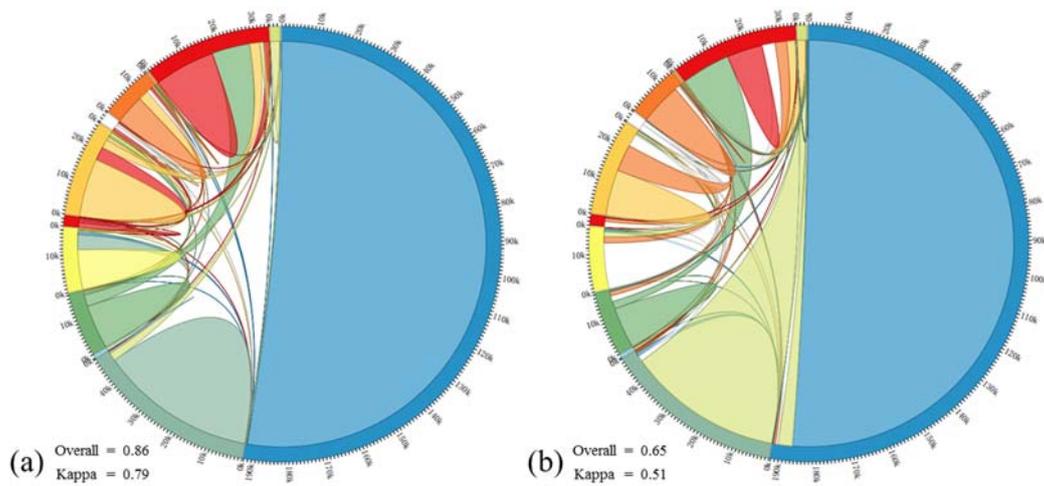


Figure 15. Chord diagram of the error matrix in Case 2: (a) between the reference and interpretation of the proposed method, (b) between the reference and interpretation of the supervised classification (the legend in the diagram is the same as in Figure 14).

The interpretation using supervised classification has an overall accuracy of 0.65 and a kappa of 0.51 with a total of seven classes identified. Here, the class in white refers to image objects with high reflectance and an elongated shape, including the unsealed roads, the sealed roads, and the banks of the reclamation ponds. These image objects cannot be classified further. Similarly, the supervised classification performs poorly in the interpretation of ponds. First, the category defined as pond means that a pond of water should be of the same kind. Second, regarding image objects independently ignores the fact that they provide context to each other.

Generally, the two cases were interpreted better using the proposed approach compared to supervised classification. In Case 1, with the proposed approach, the number of classes and the overall accuracy increased to nine and 0.22, respectively. In Case 2, the increases were five and 0.21, respectively. The reason why supervised classification performed poorly is due to intended classification-type meanings containing context that cannot be analyzed by supervised classification. The context is that the neighbors, the components, and the large regions together give each other meanings. This is an important way for humans to recognize reality. The two cases highlight the role of context for the interpretation improvement of more conceptual information. In summary, the proposed approach provides a deeper understanding of the image.

4. Conclusions

This study implemented an ontology-based remote sensing image interpretation, presenting a novel way to use context with spectral features, spatial features, and topological relationships. Two HSR images of coastal areas were interpreted by the proposed approach, with supervised classification serving as a contrast. Error matrices were computed to evaluate the results. The number of classes increased and the total accuracy improved in both regions.

In the proposed approach, ontology played a powerful role in semantics formalization, making the knowledge expert guide the interpretation directly. This provides an opportunity to analyze images like humans. There is huge potential to develop ontology-based methods in remote sensing. The core problem with knowledge-guided interpretation is the semantic gap, a problem solved by expert. One solution is to define geographic concepts using image features. Geographic concepts must be clearly defined, and at the same time, the feature extraction is crucial, especially developing feature expressions that can connect to the high-level concept semantics. Context provides the geographic concepts in the interpretation a higher level in the cognitive perspective. In the proposed approach, the results show the number of interpreted classes increased and the total accuracy improved by comparison to supervised classification. This suggests that the approach is an effective way to express context by combining the geographic concepts definition with the multi-scaled image object.

There are limitations to the proposed approach. Using the model for multi-scaled image objects to express context is based on the assumption that the context provider is a segmented image object. That means the context provider is only the region's spectrally homogeneous inside and heterogeneous outside. Like all ontology-based image interpretation methods, it depends on a specific application. These dependencies exist throughout each step. Therefore, future research points to independency and universality as good directions to explore.

Acknowledgments: This study was supported by the National Natural Science Foundation of China (Grant Nos. U1609202, 41376184 and 40976109), the National Key Research and Development Program of China (Grant No. 2016YFC1400903), and the R&D Special Fund for Public Welfare Industry (Oceanography, Grant Nos. 201005011 and 201305009).

Author Contributions: Jianyu Chen outlined the research topic, assisted with manuscript writing and coordinated the revision. Helingjie Huang implemented the method, performed the data preprocessing, and wrote the manuscript. Zhu Li was involved in data collection. Ninghua Chen and Fang Gong performed the field investigations. All authors participated in editing and revising the paper.

Conflicts of Interest: The authors declare no conflicts of interest.

References

- Blaschke, T.; Hay, G.J.; Kelly, M.; Lang, S.; Hofmann, P.; Addink, E.; Queiroz Feitosa, R.; van der Meer, F.; van der Werff, H.; van Coillie, F.; et al. Geographic object-based image analysis—Towards a new paradigm. *ISPRS J. Photogramm. Remote Sens.* **2014**, *87*, 180–191. [[CrossRef](#)] [[PubMed](#)]
- Baatz, M.; Schäpe, A. Multiresolution Segmentation: An optimization approach for high quality multi-scale image segmentation. *J. Photogramm. Remote Sens.* **2000**, *58*, 12–23.
- Chen, J.; Pan, D.; Mao, Z. Image-object detectable in multiscale analysis on high-resolution remotely sensed imagery. *Int. J. Remote Sens.* **2009**, *30*, 3585–3602. [[CrossRef](#)]
- Chen, G.; Hay, G.J.; St-Onge, B. A GEOBIA framework to estimate forest parameters from lidar transects, Quickbird imagery and machine learning: A case study in Quebec, Canada. *Int. J. Appl. Earth Obs. Geo-inf.* **2012**, *15*, 28–37. [[CrossRef](#)]
- Benz, U.C.; Hofmann, P.; Willhauck, G.; Lingenfelder, I.; Heynen, M. Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. *ISPRS J. Photogramm. Remote Sens.* **2004**, *58*, 239–258. [[CrossRef](#)]
- eCognition Professional User Guide. Available online: <http://www.definiens-imaging.com/> (accessed on 19 January 2017).
- Hajj, E.M.; Bégue, A.; Guillaume, S.; Martiné, J.F. Integrating SPOT-5 time series, crop growth modeling and expert knowledge for monitoring agricultural practices—The case of sugarcane harvest on Reunion Island. *Remote Sens. Environ.* **2011**, *113*, 2052–2061. [[CrossRef](#)]
- Forestier, G.; Puissant, A.; Gancarski, P.; Wemmert, C.; Ganarski, P. Knowledge-based Region Labeling for Remote Sensing Image Interpretation. *Comput. Environ. Urban Syst.* **2012**, *36*, 470–480. [[CrossRef](#)]
- Moller-Jensen, L. Classification of urban land cover based on expert systems, object models and texture. *Comput. Environ. Urban Syst.* **1997**, *21*, 291–302. [[CrossRef](#)]
- Lillesand, T.M.; Kiefer, R.W.; Chipman, J.W. *Remote Sensing and Image Interpretation*, 7th ed.; Wiley: Hoboken, NJ, USA, 2003.
- Gruber, T.R. A translation approach to portable ontology specifications. *Knowl. Acquis.* **1993**, *5*, 199–220. [[CrossRef](#)]
- Andres, S.; Arvor, D.; Pierkot, C. Towards an ontological approach for classifying remote sensing images. In Proceedings of the 2012 Eighth International Conference on Signal Image Technology and Internet Based Systems (SITIS), Naples, Italy, 25–29 November 2012.
- Forestier, G.; Wemmert, C.; Puissant, A. Coastal image interpretation using background knowledge and semantics. *Comput. Geosci.* **2013**, *54*, 88–96. [[CrossRef](#)]
- Luo, H.; Li, L.; Zhu, H.; Kuai, X.; Zhang, Z.; Liu, Y. Land Cover Extraction from High Resolution ZY-3 Satellite Imagery Using Ontology-Based Method. *ISPRS Int. J. Geo Inf.* **2016**, *5*, 31. [[CrossRef](#)]
- Durand, N.; Derivaux, S.; Forestier, G.; Wemmert, C.; Gancarski, P.; Boussaid, O.; Puissant, A.; Ganc, P.; Boussa, O.; Puissant, A. Ontology-based Object Recognition for Remote Sensing Image Interpretation. In Proceedings of the 19th IEEE International Conference on Tools with Artificial Intelligence (ICTAI 2007), Washington, DC, USA, 29–31 October 2007; Volume 1, pp. 472–479.
- Puissant, A.; Sheeren, D.; Durand, D. Urban ontology for semantic intergretation of muti-source images. In Proceedings of the 2nd Workshop on Ontologies for Urban Development: Conceptual Models for Practitioners, Turin, Italy, 17–18 October 2007; pp. 1–17.
- Fikes, R.; Hayes, P.; Horrocks, I. OWL-QL—A language for deductive query answering on the Semantic Web. *Web Semant. Sci. Serv. Agents World Wide Web* **2004**, *2*, 19–29. [[CrossRef](#)]
- Addink, E.A.; Van Coillie, F.M.B.; de Jong, S.M. Introduction to the GEOBIA 2010 special issue: From pixels to geographic objects in remote sensing image analysis. *Int. J. Appl. Earth Obs. Geoinf.* **2012**, *15*, 1–6. [[CrossRef](#)]
- Arvor, D.; Durieux, L.; Andrés, S.; Laporte, M.-A. Advances in Geographic Object-Based Image Analysis with ontologies: A review of main contributions and limitations from a remote sensing perspective. *ISPRS J. Photogramm. Remote Sens.* **2013**, *82*, 125–137. [[CrossRef](#)]
- Cheng, J.; Bo, Y.; Zhu, Y.; Ji, X. A novel method for assessing the segmentation quality of high-spatial resolution remote-sensing images. *Int. J. Remote Sens.* **2014**, *35*, 3816–3839. [[CrossRef](#)]

21. Schöpfer, E.; Lang, S. Object fate analysis—A virtual overlay method for the categorisation of object transition and object-based accuracy assessment. In Proceedings of the 1st International Conference on Object-Based Image Analysis, Salzburg, Austria, 4–5 July 2006.
22. Lucieer, A.; Stein, A. Existential uncertainty of spatial objects segmented from satellite sensor imagery. *IEEE Trans. Geosci. Remote Sens.* **2002**, *40*, 2518–2521. [[CrossRef](#)]
23. Navulur, K. *Multispectral Image Analysis Using the Object-Oriented Paradigm*; CRC Press: Boca Raton, FL, USA, 2006.
24. OWL Working Group OWL 2 Web Ontology Language Document Overview (Second Edition). Available online: <https://www.w3.org/TR/owl-overview/> (accessed on 6 March 2017).
25. Krötzsch, M. OWL 2 Profiles: An introduction to lightweight ontology languages. In *Reasoning Web International Summer School*; Springer: Berlin/Heidelberg, Germany, 2012; pp. 112–183.
26. REQUIEM: REsolution-Based QUery Rewriting for Expressive Models. Available online: <http://www.cs.ox.ac.uk/projects/requiem/> (accessed on 12 December 2016).
27. Pérez-Urbina, H.; Motik, B.; Horrocks, I. Tractable query answering and rewriting under description logic constraints. *J. Appl. Log.* **2010**, *8*, 186–209. [[CrossRef](#)]
28. Ji, X. Research on the Method of Accuracy Assessment of the Object-Based Classification from Remotely Sensed Data. Master's Thesis, Beijing Normal University, Beijing, China, 2012.



© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).