Operational Performance of an Automatic Preliminary Spectral Rule-Based Decision-Tree Classifier of Spaceborne Very High Resolution Optical Images

Andrea Baraldi, Tom Wassenaar, and Simon Kay

***Abstract*—In the last 20 years, the number of spaceborne very high resolution (VHR) optical imaging sensors and the use of satellite VHR optical images have continued to increase both in terms of quantity and quality of data. This has driven the need for automating quantitative analysis of spaceborne VHR optical imagery. Unfortunately, existing remote sensing image understanding systems (RS-IUSs) score poorly in operational contexts. In recent years, to overcome operational drawbacks of existing RS-IUSs, an original two-stage stratified hierarchical RS-IUS architecture has been proposed by Shackelford and Davis. More recently, an operational automatic pixel-based near-real-time four-band IKONOS-like spectral rule-based decision-tree classifier (ISRC) has been downscaled from an original seven-band Landsat-like SRC (LSRC). The following is true for ISRC: 1) It is suitable for mapping spaceborne VHR optical imagery radiometrically calibrated into top-of-atmosphere or surface reflectance values, and 2) it is eligible for use as the pixel-based preliminary classification first stage of a Shackelford and Davis two-stage stratified hierarchical RS-IUS architecture. Given the ISRC "full degree" of automation, which cannot be surpassed, and ISRC computation time, which is near real time, this paper provides a quantitative assessment of ISRC accuracy and robustness to changes in the input data set consisting of 14 multisource spaceborne images of agricultural landscapes selected across the European Union. The collected experimental results show that, first, in a dichotomous vegetation/nonvegetation classification of four synthesized VHR images at regional scale, ISRC, in comparison with LSRC, provides a vegetation detection accuracy ranging from 76% to 97%, rising to about 99% if pixels featuring a low leaf area index are not considered in the comparison. Second, in the generation of a binary vegetation mask from ten panchromatic-sharpened QuickBird-2 and IKONOS-2 images, the operational performance measurement of ISRC is superior to that of an ordinary normalized difference vegetation index thresholding technique. Finally, the second-stage automatic stratified texture-based separation of low-texture annual cropland or herbaceous range land (land cover class AC/HR) from high-texture forest or woodland (land cover class F/W) is performed in the discrete, finite, and symbolic ISRC map domain in place of the ordinary continuous varying, subsymbolic, and multichannel texture feature domain. To conclude, this paper demonstrates that the automatic ISRC is eligible for use in operational VHR satellite-based measurement systems such as those envisaged under the ongoing Global Earth Observation System of Systems (GEOSS) and Global Monitoring for the Environment and Security (GMES) international programs.**

***Index Terms*—Image classification, image understanding system (IUS), inductive and deductive inference, prior spectral knowledge, radiometric calibration, texture analysis.**

I. Introduction

THE POTENTIAL of Earth observation (EO) from space for the monitoring of the Earth's environment and for the detection of its temporal variations at local, regional, continental, and global geographic extents is well known by user communities involved with urban growth assessment and planning, intelligence/surveillance applications for national security and defense purposes, ecosystem management, watershed protection, water balance calculations, risk management, and global change [1].

The expected impact of remote sensing (RS) imagery upon the general public has increased after the recent announcement by the Group on EOs (GEO) that scientists and decision makers around the world will soon have unrestricted access at no charge to the Landsat archive, which is the world's most extensive collection of continuously acquired RS spaceborne imagery [2]. This news followed the decision by the China–Brazil Earth Resources Satellite to distribute its images free of charge starting from 2007. In turn, the European Union (EU) announced a free access data policy for the Sentinel-2/-3 satellites whose launch is scheduled starting from 2012.

While cost-free access to large-scale low-spatial-resolution (low SR or LR; above 40 m) and medium-SR (MR; from 40 to 15 m) spaceborne image databases is becoming a reality, the demand for high-SR (HR; between 15 and 5 m) and very-high-SR (VHR; below 5 m) satellite images has continued to increase in terms of both quantity and quality of data. For example, the U.S. Government spent $56 million in 1991 and $200 million in 2003 for the purchase of VHR satellite images [3]. This increasing request for VHR spaceborne imagery has boosted the rapid growth of the commercial VHR satellite industry. For example, by 2012, current HR and VHR spaceborne optical imaging

Manuscript received July 29, 2009; revised January 1, 2010. Date of publication May 24, 2010; date of current version August 25, 2010.

A. Baraldi was with the Spatial Data Infrastructures Unit, Institute for Environment and Sustainability, Joint Research Center, European Commission, 21020 Ispra, Italy. He is now with the Baraldi Consultancy in Remote Sensing, 40129 Bologna, Italy (e-mail: andrea6311@gmail.com).

T. Wassenaar was with the Monitoring Agricultural Resources Unit, Institute for the Protection and Security of the Citizen, Joint Research Center, European Commission, 21020 Ispra, Italy. He is now with the Environmental Risks of Recycling Unit, CIRAD, 97408 Saint-Denis Cedex 09, France.

S. Kay is with the Monitoring Agricultural Resources Unit, Institute for the Protection and Security of the Citizen, Joint Research Center, European Commission, 21020 Ispra, Italy.

Digital Object Identifier 10.1109/TGRS.2010.2046741

sensors such as GeoEye-1, QuickBird-2 (QB2), IKONOS-2 (IK2), OrbView-3, KOMPSAT-2, RapidEye, FORMOSAT-2, and ALOS AVNIR-2 will be joined by planned VHR spaceborne imaging missions such as PLEIADES-1/-2, WorldView-2, and Astrium SPOT-6/-7.

These multiple drivers make urgent the need to develop operational satellite-based measurement systems that are suitable for automating the quantitative analysis of large-scale spaceborne multisource multiresolution image databases. This ambitious goal is envisaged under ongoing international programs such as the following:

1) The Global EO System of Systems (GEOSS), conceived by GEO [4], [5], which requires harmonization and interoperability of EO data and derived information products generated from a variety of sources at all scales—global, regional, and local;
2) The Global Monitoring for the Environment and Security (GMES), which is an initiative led by the EU in partnership with the European Space Agency, whose aim is to guarantee the sustainability of integrated operational services for EU security and environmental monitoring based on EO data from multiple sources (satellite, airborne, and *in situ*) and synergistic data products [6], [7].

Well known in existing literature as (2-D) object-based image analysis (OBIA) [8], the two-stage segment-based RS image understanding system (RS-IUS) architecture is currently considered as the state of the art in spaceborne image analysis [9]–[11]. In spite of the increasing popularity of commercial two-stage segment-based RS-IUSs, the automatic or semiautomatic transformation of large-scale spaceborne LR to VHR image databases into information still remains far more problematic than might be reasonably expected [12], [13]. This is tantamount to saying that existing RS-IUSs score poorly in operational performance encompassing the following [14]: 1) ease of use (degree of automation); 2) effectiveness (e.g., classification accuracy); 3) efficiency (e.g., computation time and memory occupation); 4) economy (costs); 5) robustness to changes in the input data set; 6) robustness to changes in the input parameters; 7) maintainability/scalability/reusability to keep up with users' changing needs; and 8) timeliness (defined as the time span between data acquisition and product delivery to the end user; it increases monotonically with manpower, e.g., required to collect scene-specific training samples).

A measurement of the operational performance of existing RS-IUSs scores low due to a combination of factors. *In primis*, the increasing rate of collection of spaceborne EO data of enhanced spatial, spectral, and temporal resolution outpaces the capabilities of both manual inspection and semiautomatic inductive-labeled (supervised) data learning algorithms [15]. This means that, to date, the cost, timeliness, quality, and availability of adequate reference (training/testing) data sets derived from field sites, existing maps, and tabular data can be considered the most limiting factors on inductive-labeled (supervised) RS data product generation and validation [1].

In addition, (3-D) object recognition is increasingly difficult in (2-D) imagery featuring the following: 1) finer SR, i.e., enhanced spatial (contextual) information and 2) coarser spectral resolution, i.e., less discriminative color (noncontextual chromatic and achromatic [16]) properties. The result is that the complexity of (3-D) object detection in EO optical imagery has increased monotonically during the last 20 years, while the transition from early LR and MR (e.g., seven-band 30-m Landsat sensor series) to recent HR and VHR (e.g., four-band 4-m-resolution IKONOS) spaceborne imaging sensors has been accomplished. For example, in spaceborne HR and VHR optical imagery, geometric morphological and texture properties together with spatial topological (e.g., adjacency and inclusion) and nontopological (e.g., distance and orientation) relationships become especially important for the recognition of visible man-made objects such as buildings, roads, and agricultural fields [17]–[21], [35], [39], [47].

To date, the operational drawbacks of existing RS-IUSs are well recognized by significant sections of RS literature [8], [13], [17], [22]. For example, starting from the four levels of understanding of an information processing device proposed by Marr, namely, the level of computational theory (architecture), knowledge/information representation, algorithms, and hardware and code implementations, Marr states that the lynchpin of success in attempting to solve the computer vision problem is addressing the computational theory rather than algorithms or implementations. In line with this intuition by Marr, i.e., starting from the customary distinction between a model and the algorithm used to identify it [23], [24], one of the present authors identified an original hybrid two-stage stratified hierarchical RS-IUS architecture in several RS-IUS implementations proposed by Shackelford and Davis in recent years [19]–[21]. More recently [22], [25], an operational automatic pixel-based near-real-time four-band IKONOS-like spectral rule-based decision-tree classifier (ISRC) has been downscaled from an original seven-band Landsat-like SRC (LSRC) [26], [27]. The following is true for ISRC: 1) It is "fully automated," i.e., it requires neither user-defined parameters nor reference data samples to run [28]; 2) it is suitable for mapping four-band spaceborne VHR imagery consisting of three visible bands and one near-infra-red (NIR) channel, required to be radiometrically calibrated into top-of-atmosphere (TOA) reflectance or surface reflectance values; and 3) it is eligible for use as the pixel-based preliminary classification first stage of a Shackelford and Davis two-stage stratified hierarchical RS-IUS architecture.

Due to its spectral resolution that is inferior to the seven-band LSRC, the four-band ISRC is theoretically expected to be less effective than LSRC in the separation of vegetation from bare soils due to the loss of the Landsat-like medium-IR (MIR) channels [22] and in the separation of snow from clouds and light-toned bare soils due to the loss of the MIR and thermal IR (TIR) channels [22], [29].

Given the ISRC "full degree" of automation, which cannot be surpassed, and ISRC computation time, which is near real time (e.g., a standard desktop computer requires 2–5 min to map a space image with LSRC and its downscaled versions such as ISRC [22], [25]), the objective of this paper is to provide an operational performance measurement of the following: 1) ISRC accuracy and robustness to changes in the input data set acquired across time, space, and sensors at continental scale and 2) the effectiveness of the two-stage stratified hierarchical

RS-IUS model employing ISRC as its preliminary classification first stage. To achieve these objectives, a testing data set was selected, comprising 14 spaceborne multisource real or appropriate synthetic four-band images depicting a variety of agricultural landscapes met across the EU in the spring season (from early April to late June). Consequently, this paper provides a so-called stage 1 validation of ISRC accuracy in the target spring period and a so-called stage 2 validation of ISRC accuracy over a significant range of geographic conditions and imaging sensors [1, p. 43].

The rest of this paper is organized as follows. Previous works related to VHR image understanding are reviewed in Section II. Section III describes the study areas and testing images. In Section IV, the ISRC experimental session is presented and discussed. The summary and conclusion are reported in Section V.

II. Related Works

Marr considered a computer vision system as an example of an information processing device to be understood at the following four levels of analysis [23].

1) *Computational theory:* It identifies the primary objectives of the artificial vision system and presents an artificial vision system model (architecture) required to address the complexity of the vision task in terms of the following [14]: input and output data flows, control flows, and the hierarchical sequence of data processing steps. Marr stressed that the lynchpin of success in attempting to solve the computer vision problem is to address the computational theory rather than algorithms or implementations [23], [24]. In other words, if the vision device architecture is inadequate, even sophisticated algorithms can produce low-quality outputs. On the contrary, improvement in the vision system architecture may achieve twice the benefit with half the effort (which is an adaptation of the original words by Wang: see this text in the succeeding discussion).
2) *Knowledge/information representation [23], [24], [30]:* In Wang's words, "good computer information processing depends upon a good representation method that is well-suited to the nature of the information to be processed. However, this point is not always given the deserved attention in developing new processing techniques and improving the processing quality and efficiency. Quite often, efforts are mainly made on algorithms. If knowledge representation is poor, even sophisticated algorithms can produce inferior outputs. On the contrary, improvement in representation might achieve twice the benefit with half the effort" [30].
3) Algorithm design to manipulate information representations, namely, to transform the structure of the data, to transform the information contained in the data, or to generate the new information from the data [14], [23], [24], [31]. Structured system design can be identified as "everything but code" [14].
4) Implementation to turn what was produced during the design into hardware and code [14].

In terms of computational theory, any imaging sensor projects a (3-D) scene onto a (2-D) image. Thus, the main role of a biological or artificial visual system is to back-project the information in the image domain to that in the scene domain. In greater detail, the goal of a visual system is to provide plausible (multiple) symbolic description(s) of the scene depicted in an image by finding associations between subsymbolic (2-D) image features with symbolic (3-D) objects in the scene (e.g., buildings and roads) [17], [32]. Subsymbolic (2-D) image features are either points or regions or, vice versa, region boundaries, i.e., edges provided with no semantic meaning [17]. In literature, (2-D) image regions are also called segments, (2-D) objects, patches, parcels, or blobs [33]–[35]. There is a well-known information gap between symbolic information in the (3-D) scene and subsymbolic information in the (2-D) image (e.g., due to dimensionality reduction and occlusion phenomena). This is called the *intrinsic insufficiency* of image features [17]. It means that the problem of image understanding is inherently ill posed and consequently very difficult to solve [17].

Based on the aforementioned customary distinction between a model and its implementation [14], [23], [24], [31], this section compares two alternative RS-IUS models, namely, the two-stage segment-based RS-IUS architecture, currently considered the state of the art in commercial RS image analysis software toolboxes [9]–[11], and the Shackelford and Davis two-stage stratified hierarchical RS-IUS architecture recently presented in the RS literature [19]–[21]. Next, this section revises the original operational automatic seven-band LSRC and its downscaled four-band ISRC version. The latter is eligible for use as the pixel-based preliminary classification first stage of an operational two-stage stratified hierarchical RS-IUS implementation that is suitable for mapping spaceborne VHR optical imagery [22], [25].

A. RS-IUS Architectures

1) Two-Stage Segment-Based RS-IUS Architecture: Known in existing literature as OBIA [8], the two-stage segment-based RS-IUS architecture has recently gained a noteworthy popularity in commercial RS image analysis software toolboxes [9]. Main functional properties [9] and well-known limitations [8], [13], [17], [22] of the two-stage segment-based RS-IUS model are summarized in the succeeding discussion.

a) First-stage unlabeled data-driven (bottom-up) inherently ill-posed driven-without-knowledge segmentation:

1) Although acknowledged by a reasonable portion of the existing literature [8], [17], [36], [37], the inherent ill-posedness of image segmentation/edge detection, the latter being the dual problem of the former, is often forgotten by a large segment of the RS community. This may explain why, although no "best" segmentation/edge detection approach exists, literally dozens of "novel" (supposedly better!) segmentation/edge detection algorithms are published each year [12]. Due to its inherent ill-posedness, the problem of image segmentation is very difficult to solve. In RS common practice, this implies the following.
   a) Any segmentation algorithm is affected by the so-called *artificial insufficiency* of image segments. This is tantamount to saying that, in real image

segmentation problems (other than toy problems), it is inevitable that erroneous segments are detected while genuine segments are omitted [17, p. 18]. The concept of *artificial insufficiency* of image segments is directly linked to the well-known *uncertainty principle*; according to it, for any contextual (neighborhood) property, we cannot simultaneously measure that property while obtaining accurate localization [13], [38].

b) To make the ill-posed driven-without-knowledge segmentation problem better posed, segmentation algorithms require system parameters to be user defined based on heuristic criteria [11]. As a consequence, segmentation algorithms are the following.
   i) Poorly posed semiautomatic rather than well-posed automatic, where the latter is representing the best case in operational terms with respect to ease of use.
   ii) Not robust to changes in the input data set. Typically, a driven-without-knowledge segmentation algorithm requires a novel set of user-defined parameters per input scene.
   iii) Difficult to use. User-defined parameters are not provided with a clear unequivocal physical meaning and/or a known range of variation. In addition, parameter values are selected based on empirical criteria.

c) To summarize, due to their intrinsic ill-posedness, RS image segmentation algorithms score low in operational performance in terms of the following: 1) degree of automation (ease of use); 2) accuracy; and 3) robustness to changes in the input data set.

2) Any driven-without-knowledge segmentation algorithm employs as input subsymbolic (unlabeled and asemantic) pixels and it provides, as output, subsymbolic informational primitives featuring one single spatial type, namely, subsymbolic (2-D) regions (segments). Locational, photometric, geometric, and shape properties of image segments and their spatial relationships can be described in a segment description table [17], [39].

3) In RS common practice, an ill-posed driven-without-knowledge segmentation first stage is often implemented as a multiscale (hierarchical) iterative segmentation algorithm [9]. As the output, a hierarchical segmentation algorithm generates multiscale segmentation solutions in the hope that the target image will appear correctly segmented at some scale. Unfortunately, quantitative multiscale assessment of segmentation quality indices requires ground truth data at each scale that are impossible or impractical to obtain in RS common practice [13]. Therefore, the "best" segmentation map must be selected by the user on an *a posteriori* basis from the available set of multiscale segmentation solutions according to heuristic, subjective, and/or qualitative criteria that are analogous to those employed in the selection of prior segmentation parameters. To conclude, exploitation of a hierarchical segmentation algorithm does not make the driven-without-knowledge segmentation first stage easier to use. In addition, hierarchical segmentation algorithms are computationally intensive and require large memory occupation.

b) Second-stage segment-based classification modules:

1) They can be implemented either top-down (model driven), such as deductive decision-tree classifiers exclusively based on prior knowledge of the (3-D) world, or bottom-up (data driven), such as inductive labeled-data (supervised-data) learning classifiers (e.g., artificial neural networks) [15], [40], [41] or inductive unlabeled-data (unsupervised-data) learning distributed systems (e.g., unlabeled data clustering algorithms) [15]. In practice, under the guise of "flexibility," two-stage segment-based RS-IUSs provide RS experts and practitioners with overly complicated machine learning options to choose from based on heuristic criteria. This is tantamount to saying that two-stage segment-based RS-IUSs are difficult to use and require expert users [8].

2) According to Section I, the current rate of the collection of RS data outpaces the manual capability for gathering ground truth data. In many RS data application problems, the costs, quality, and availability of adequate training labeled (reference) samples are the most limiting factors on the exploitation of inductive labeled-data (supervised-data) learning algorithms. This means that the universal problem of interfacing numerical (quantitative) computation with symbolic (qualitative) computation remains difficult to solve with two-stage segment-based RS-IUSs [17].

3) Due to poor knowledge/information representation of image features, consisting of intrinsically unreliable subsymbolic regions exclusively, employed as input by the second-stage segment-based classification modules, the two-stage segment-based RS-IUSs lack flexibility and reasoning capability (which increases with the accumulation of evidence). For example, the computational efficiency of commercial two-stage segment-based RS-IUSs, such as [9], becomes extremely low when image information should be dealt with at pixel rather than segment level, e.g., when a simple spectral decision rule should be applied pixelwise, i.e., when dealing with segments consisting of a single pixel.

4) Second-stage supervised classification modules provide the unlabeled data-driven segmentation first stage with no feedback mechanism to recover from the artificial insufficiency of image regions.

To conclude, there is still a lack of consensus and research on the conceptual foundation of OBIA, i.e., on the relationship between inherently ill-posed subsymbolic (asemantic) (2-D) image segments (regions) and symbolic (semantic) (3-D) landscape objects (e.g., *forest*, *agricultural field*, etc.) [8], [13], [22].

2) Two-Stage Stratified Hierarchical RS-IUS Architecture: In recent years, Shackelford and Davis have presented several implementations of an RS-IUS that is suitable for mapping 1-m-resolution panchromatic (PAN)-sharpened multispectral (MS) IKONOS imagery of urban areas (see Fig. 1) [19]–[21]. Based on the customary distinction between a model and the algorithm used to identify it made by computational theory [14], [23], [24], [31], one of the present authors revealed the presence of a novel two-stage stratified hierarchical RS-IUS architecture in the original Shackelford and Davis RS-IUS implementations [22], [25], [27], [42]. The main

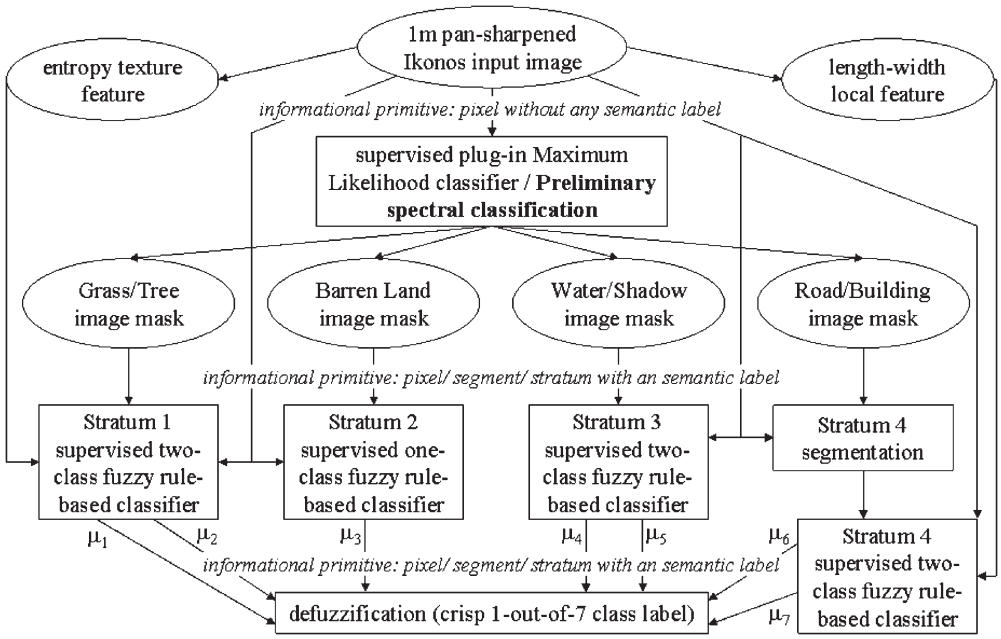


Fig. 1. Two-stage stratified hierarchical RS-IUS implementation proposed by Shackelford and Davis [19]–[21]. As the output, the application-independent pixel-based preliminary classification first stage provides simultaneously symbolic pixels in symbolic segments (defined as connected sets of labeled pixels featuring the same label) in symbolic strata (defined as image-wide sets of labeled pixels featuring the same label). All provided with a semantic meaning which is well-understood by a user. The second stage consists of stratified land cover class-specific context-sensitive feature extractors and fuzzy rule-based classifiers.

functional properties of the two-stage stratified hierarchical RS-IUS architecture and the main limitation of the Shackelford and Davis RS-IUS implementations are summarized in the succeeding discussion (see Fig. 1).

a) Preliminary pixel-based model-driven (top-down) classification first stage:

1) A preliminary pixel-based model-driven (top-down) classification first stage provides a *primal sketch* or *preliminary map* in the Marr sense [23]. A primal sketch is theoretically defined by Marr as the (2-D) image information describing the "geometrical distribution and organization of" either color or "intensity changes" (edges) (refer to [23, p. 37]). To underline the customary distinction between a model and the algorithm used to identify it [14], [23], [24], [31], it is worth mentioning that the primal sketch is computed by Marr in two computation steps that have nothing to do with the pixel-based model-driven classifier adopted by Shackelford and Davis. In particular, the Marr preliminary map detection first and second steps are the following [23].
   a) Marr's first step: either color-, intensity-, or texture-based edges are detected using context-sensitive Laplacian filters at different scales.
   b) Marr's second step: domain-independent general-purpose (e.g., Gestalt) grouping methods pursue an edge-preserving aggregation (segmentation) of pixels into perceptually "homogeneous" (2-D) objects (termed "blobs, edge segments, and groups").
2) In the Shackelford and Davis RS-IUS implementations, the preliminary pixel-based model-driven (top-down) classification first stage is implemented as a supervised pixel-based plug-in (i.e., noniterative and one pass) maximum likelihood (ML) classifier. Like any plug-in ML classifier, it maps each pixel onto a discrete and finite set of semantic (symbolic) labels. However, unlike traditional ML classification of RS data where semantic labels identify target land cover classes, in the Shackelford and Davis ML implementation, the labels identify the land cover class sets [19], [20], i.e., "combine" sets of primitive land cover classes. In greater detail, each land cover class set may comprise the following: 1) a single land cover class when its spectral overlap with the other land cover classes belonging to the classification scheme is negligible (e.g., *'barren land'*, refer to Fig. 1) [43] or 2) two or more land cover classes whose interclass spectral overlap is statistically significant (e.g., *'either water or shadow'*, refer to Fig. 1). To summarize, in the Shackelford and Davis ML implementation, the interset degree of spectral overlap is negligible, while the degree of intraset spectral overlap is statistically relevant. This is tantamount to saying that the land cover class sets are mutually exclusive. As a consequence, the first-stage pixel-based plug-in ML classifier implemented by Shackelford and Davis is well posed, i.e., crisp, unequivocal, and reliable. This means that the primary cause of the *artificial insufficiency* of image features (i.e., points and regions or, vice versa, edges) in two-stage RS-IUS instantiations can be removed by replacing an intrinsically ill-posed unreliable segmentation first stage with a well-posed pixel-based model-driven preliminary classifier [17].[1]

[1]This substitution would turn into RS common practice the popular statement ascribed to Archimedes, the famous inventor, mathematician, and physicist of ancient Syracuse, Sicily (287 BC - 212 BC), who is said to have remarked about the lever: Give me a place to stand on, and I will move the Earth. This statement is consistent with the biblical teaching of the parable of the two builders, where a wise man is described as building his house on the rock, rather than on sand (Matthew 7:24–27). By analogy, a well-posed preliminary spectral-based per-pixel classifier would be the rock to stand on to build operational RS-IUSs.

3) Given the structural content of urban scenes depicted in spaceborne VHR imagery, the following five class sets are identified by Shackelford and Davis (see Fig. 1): 1) *'either grass or tree'*; 2) *'either road or building or impervious surface'*; 3) *'either water or shadow'*; 4) *'barren land'*; and 5) *'others'* (outliers, e.g., clouds). In line with the Congalton requirements, this classification scheme is mutually exclusive and totally exhaustive [43, p. 12].
4) In [22], [25], [27], and [42], land cover class sets detected as output by the pixel-based preliminary classification first stage are called *spectral categories*, *spectral types*, *spectral layers*, *spectral strata*, or *spectral-based semiconcepts*. A spectral-based semiconcept is a *semantic conjecture (e.g., 'vegetation' as 'either grass or tree') based solely on the per-pixel (noncontextual) color (spectral, i.e., chromatic and achromatic* [16]*) properties.*
5) The semantic stratification provided by the preliminary classification first stage acts as a *focus of visual attention*[32], [33], [44]: It identifies places of interest in the image that should be processed with higher accuracy (and increasing computation time) at the second stage [24]. In other words, semantic stratification identifies symbolic candidate areas of the image (e.g., stratum *'vegetation'*) with a high probability of a target (2-D) region type (e.g., a *forest* region) being present [17].
6) It is noteworthy that, while (3-D) land cover classes (e.g., *forest*) are provided with a high-level semantic meaning but are difficult to detect automatically in (2-D) images, the following is true for spectral-based semiconcepts (e.g., *'vegetation'*): 1) They can be detected automatically in MS imagery exclusively based on prior spectral knowledge [22], [26], [27], and 2) they are provided with a semantic meaning that is not superior (i.e., equal or inferior) to that of the target (3-D) land cover classes but superior to zero of the traditional subsymbolic image information primitives, such as subsymbolic regions and unlabeled data clusters [15]. Therefore, spectral-based semiconcepts are eligible for filling in the well-known information gap existing between symbolic information in the (3-D) scene and subsymbolic information in the (2-D) image [17].
7) As the output, a preliminary pixel-based model-driven classification first stage provides symbolic information primitives featuring three spatial types, namely, symbolic pixels in symbolic segments (defined as connected sets of labeled pixels featuring the same label) in symbolic strata (defined as image-wide sets of labeled pixels featuring the same label). These three spatial types are not alternative but coexist and can be selected according to specific needs of the second-stage battery of application-specific satellite-based measurement systems (refer to Section II-A2b).
8) From the computational point of view, since it is pixel based and prior model driven, the first-stage preliminary classifier is context insensitive and one pass (noniterative). Therefore, it is computationally efficient.
9) Since the first-stage preliminary classifier is pixel based, it works at the sensor SR. In practice, it is sensor resolution independent.

b) Second-stage hierarchy of stratified (3-D) object model-specific classification modules for spatial reasoning:

1) The battery of stratified hierarchical class-specific classification modules incorporates the "stratified" or "layered" approach that is typical of decision trees [19]–[21], [45]. This is tantamount to saying that this hierarchy enforces the well-known divide-and-conquer (*dividi et impera*) problem-solving approach [45].
2) The idea of stratification is well known in statistics. For example, in stratified sampling, the sampling frame is divided into nonoverlapping groups or strata, e.g., geographical areas. A sample is taken from each stratum, and when this sample is a simple random sample, it is referred to as stratified random sampling. A possible disadvantage of stratification is that identifying the appropriate strata may be difficult. *The advantage is that stratification will always achieve greater precision provided that the strata have been chosen so that members of the same stratum are as similar as possible in respect of the characteristic of interest* [46]. For example, any given ill-posed segmentation algorithm, typically applied image-wide on a "dumb" (driven without knowledge) basis (as in two-stage segment-based RS-IUSs), achieves greater precision (because better conditioned) when it is run separately on mutually exclusive image strata where the image segmentation subproblem becomes easier to solve. This principle is exploited in Fig. 1, where a second-stage better-posed stratified segmentation algorithm is run to detect shape properties of (2-D) regions eligible for assignment to man-made (3-D) objects, such as *buildings*, *roads*, and *agricultural fields*, whose geometric attributes, morphological properties, and spatial relationships are especially important for their recognition.
3) The battery of stratified hierarchical land cover class-specific classification modules comprises the following (see Fig. 1).
   a) Stratified context-sensitive (e.g., texture, geometric, morphological, etc.) (2-D) image feature extraction modules. Contextual information is computationally demanding to extract. To increase computational efficiency of neighboring information extraction, the strategy is twofold.
      i) Since color information is dealt with by the first-stage spectral-based preliminary classifier providing, as output, semantic strata (i.e., full information is extracted from color data by the first-stage pixel-based preliminary classifier), geometric, morphological, and texture information can be generated at a (stratified) second stage from a one-band brightness (intensity and achromatic) image generated as a linear combination of MS channels [39]. Thus, in second-stage stratified contextual feature extraction modules, a one-band brightness image can replace, as input, the MS image.
      ii) In Fig. 1, contextual information is selectively computed on a stratified land-cover class-specific basis if and only if this contextual information is required

by the control knowledge guiding the reasoning processes [17]. In other words, only if considered useful is contextual information computed selectively (e.g., texture is selected for *forest* detection, while shape properties are adopted for *building* identification) within image areas (symbolic strata, e.g., *'either bare soil or built up'*) that are candidates for the presence of (2-D) regions belonging to a target (3-D) land cover class (e.g., *house*). It is noteworthy that, to date, the aforementioned spatial reasoning built on driven-by-knowledge hierarchical stratified selection of contextual information is provided with a relevant degree of novelty. For example, in recent RS literature, dozens of inductive-labeled (supervised) data learning classification black boxes (like artificial neural networks) employ, as input, a driven-without-knowledge (i.e., "blind" and unconditional) flat (i.e., nonhierarchical) combination of pixel-based (i.e., chromatic and achromatic) information with contextual properties at all pixels [76]–[79].

b) Stratified land cover class-specific fuzzy rule-based classification modules, namely, semantic or concept nets based on prior knowledge of the world, called *the world model*, consisting of (3-D) object models [17], [18], [47]. In semantic nets, nodes represent concepts, i.e., classes of (3-D) objects in the world, while edges represent relations between nodes. Typical interobject relations are PART-OF, A-KIND-OF, topological spatial relations (e.g., adjacency and inclusion) [17], [18], temporal transitions [47], etc. [24]. In these (3-D) object model-specific classification modules, a convergence-of-evidence decision mechanism is enforced [17], [39]. Accumulation of the evidence decreases the total amount of effort spent in the search for image interpretation solutions and increases the reliability of the image analysis [17].

4) At the second stage, symbolic knowledge/information representation comprises the following: symbolic pixels in symbolic segments in symbolic strata, semantic nets, and (3-D) object models belonging to the world model. Symbolic knowledge representation makes the RS-IUS easy to interact with by an application developer or expert photointerpreter who is naturally familiar with symbolic reasoning.

To summarize, in a Shackelford and Davis two-stage stratified hierarchical RS-IUS model, semantic stratification provided by the preliminary spectral-based classification first stage guarantees, at the second stage, a seamless (automatic) interface between numerical (quantitative) computation of contextual image features and symbolic (qualitative) computation. The development of such an interface is an open problem that is difficult to solve with existing commercial RS-IUSs [9], [11].

The main drawback of the two-stage stratified hierarchical RS-IUS implementations proposed by Shackelford and Davis is their need for supervised training data at every hierarchical stage [19]–[21].

TABLE I
OPTICAL SPECTRUM AND ATMOSPHERIC WINDOWS [29, p. 34]

Atmospheric window	Spectral region (μm)	Electromagnetic spectrum in optical wavelengths (μm)	Spectral region conventional name
1	0.3-1.3	Violet, V: 0.35-0.45. Blue, B: 0.45-0.50. Green, G: 0.50-0.57. Yellow, Y: 0.57-0.60. Orange, O: 0.60-0.65. Red, R: 0.65-0.72.	Visible
		0.72-1.3	Near IR (NIR)
2	1.5-1.8	1.3-3.0 (actually, 1.3-7.0)	Middle IR (MIR)
3	2.0-2.6		
4	3.0-3.6		
5	4.2-5.0		
6	7.0-15.0	7.0-15.0 (excluding TIR) 8.0-12.0	Far IR (FIR) Thermal IR (TIR)

To reduce to zero the need for supervised training data of the plug-in ML classifier implemented by Shackelford and Davis as the preliminary spectral classification first stage, the fully automated ISRC system can be adopted instead (refer to Section II-B in the following).

B. Seven-Band LSRC and Its Downscaled Four-Band ISRC Version

In [22], ISRC was proposed as a downscaled implementation of the original LSRC system presented in [26] and [27]. ISRC and LSRC are summarized in the following to make this paper self-contained.

Based exclusively on spectral prior knowledge, the operational automatic LSRC system is a one-pass (noniterative) classifier that is not adaptive to input data and that requires neither user-defined parameters nor training data samples to run. Therefore, it is called "fully automated" [28]. The LSRC base of spectral knowledge is acquired offline, i.e., before the MS image analysis takes place. It consists of a reference dictionary of endmember spectra in TOA reflectance (TOARF) values or surface reflectance ρ values, with the latter being an ideal (atmospheric noise free) case of the former, i.e., TOARF $\supseteq \rho$ [22], [27], [42]. This means the following.

1) LSRC may benefit from an inherently ill-posed atmospheric correction preprocessing stage, where user supervision becomes necessary to make the atmospheric correction problem better posed, but it does not require this. This increases the degree of automation of LSRC. On the contrary, ill-posed atmospheric correction preprocessing is considered mandatory by alternative model-driven RS image classification approaches whose prior knowledge base consists of reference surface reflectance spectra exclusively, such as the atmospheric correction for satellite imagery (ATCOR3) [48].
2) LSRC is in line with the Quality Assurance Framework for EO (QA4EO) initiative, led by the Committee of EOs Working Group on Calibration and Validation in the context of the GEOSS program [49], where an appropriate coordinated program of calibration and validation (Cal/Val) activities throughout all stages of a spaceborne

TABLE II
SPECTRAL RESOLUTION OF THE LANDSAT-4/-5 TM AND LANDSAT-7 ETM+ SENSORS IN COMPARISON WITH SPACEBORNE HR AND VHR OPTICAL IMAGING SENSORS

Landsat-4/-5 TM and Landsat-7 ETM+ (MS: 30 m SR, PAN^: 15 m)		IKONOS-2 (MS: 4 m SR, PAN: 1 m SR)		QuickBird-2 (MS: 2.44 m SR, PAN: 0.61 m SR)		AVNIR-2 (MS: 10 m SR)		RapidEye (MS: 5 m SR)	
Band	Spectral region (μm)	Band	Spectral region (μm)[1]	Band	Spectral region (μm)[2]	Band	Spectral region (μm)[3]	Band	Spectral region (μm)[4]
1 (B)	0.45-0.52	1	0.445-0.516	1	0.445-0.510	1	0.42-0.50	1	0.440-0.510
2 (G)	0.52-0.60	2	0.506-0.595	2	0.500-0.595	2	0.52-0.60	2	0.520-0.590
3 (R)	0.63-0.69	3	0.632-0.698	3	0.620-0.690	3	0.61-0.69	3	0.630-0.685
-	-	-	-	-	-	-	-	4 (Red Edge)	0.690-0.730
4 (NIR)	0.76-0.90	4	0.757-0.853	4	0.755-0.875	4	0.76-0.89	5	0.760-0.850
5 (MIR1)	1.55-1.75	-	-	-	-	-	-	-	-
7 (MIR2)	2.08-2.35	-	-	-	-	-	-	-	-
6 (TIR)	10.4-12.5	-	-	-	-	-	-	-	-

^: Landsat-7 ETM+ exclusively.
1: http://en.wikipedia.org/wiki/IKONOS.
2: http://amelia.db.erau.edu/nasacds/200606Disc1/research/20060019230_2006011209.pdf
3: http://www.eorc.jaxa.jp/ALOS/about/avnir2.htm
4: http://www.rapideye.de

TABLE III
TESTING RS IMAGE DESCRIPTION TABLE. QB2: QUICKBIRD-2. IK2: IKONOS-2. TM5: LANDSAT-5 TM

Site identifier	Site name	Country	Sensor	Imaging date	Sun elev.	Sun azim.	Viewing elev.	Viewing azim.	Scene particularity
1	ZP13	Poland	QB2	30/04/2005	52	170	61	281	Presence of unleaved deciduous forest
2	WM04	Poland	QB2	30/05/2005	57	156	57	105	-
3	HEAN	Sweden	QB2	08/05/2004	47	160	67	101	-
4	HERV	Belgium	QB2	01/04/2005	43	165	71	267	Presence of small cumulus clouds
5	CAMP	Italy	QB2	13/06/2004	68	140	66	271	Presence of cumulus clouds
6	ALMO	Spain	QB2	17/05/2004	63	140	74	123	
7	FLOR	Greece	IK2	30/04/2005	60	145	55	144	
8	PICK	UK	IK2	08/06/2004	60	164	69	356	Presence of thin haze and thick stratus clouds
9	TUAM	Ireland	IK2	01/06/2006	58	164	72	3	Presence of peat exploitation areas
10	AUME	Germany	IK2	07/06/2006	61	157	52	187	Presence of light haze
11	CAMP	Italy	TM5	20/07/2004	60	127			Presence of small cumulus clouds
12	CHAN	France	TM5	25/05/2004	57	140			
13	MURO	France	TM5	12/04/2004	48	144			
14	VRCH	Czech republic	TM5	29/04/2004	50	147			Presence de of thin stratus clouds

mission, from sensor build to end of life, is considered mandatory [49].

As the input, LSRC requires a seven-band Landsat-like image, comprised of a visible blue (B), a visible green (G), a visible red (R), a NIR, two MIRs (MIR1 and MIR2), and a TIR channel in agreement with Tables I and II, radiometrically calibrated into TOARF or ρ values [22], [42], which makes this input data set well behaved and well understood [50]. These conditions are considered *necessary, although not sufficient, requirements for input data to allow automation of a data processing system* [22], [42].

As the output, LSRC generates a preliminary classification map consisting of 46 spectral categories belonging to six parent spectral categories (supercategories), which are listed in the following (according to their order of detection and to be compared with spectral categories detected by the plug-in ML preliminary classification first stage adopted by Shackelford and Davis, refer to Section II-A2 earlier): 1) *'cloud'* (CL); 2) *'either snow or ice'* (SN); 3) *'either water or shadow'* (WASH); 4) *'vegetation'* (V); 5) *'either bare soil or built-up'* (BB); and 6) *'outliers'* [26], [27].

It is noteworthy that the aforementioned spectral-based semi-concepts are consistent with the initial dichotomous layers of well-known hierarchical RS data classification taxonomies such as the Coordination of Information on the Environment (CORINE) [51], the U.S. Geological Survey (USGS) classification hierarchy [52], and the Food and Agriculture Organization of the United Nations (FAO) Land Cover Classification System (LCCS) [53], [54].

Starting from [26], where enough information is provided to the reader for the LSRC implementation to be reproduced, the downscaled four-band ISRC system is generated by removing the excess (MIR1, MIR2, and TIR) spectral channels from the seven-band LSRC rule set while enforcing equivalences between the remaining Landsat-4/-5 Thematic Mapper (TM) and Landsat-7 Enhanced TM Plus (ETM+) bands (identified as "master" channels) with the VHR channels (identified as "slave" channels) [22] (refer to Table II). Worthy of note is that the approximation of the target Landsat band-specific sensitivity curve with the slave VHR sensor channels must fall within the band-specific range of spectral change assigned by LSRC to each spectral category, i.e., this spectral sensitivity curve approximation must fall below the within-spectral category variance. In theoretical terms, the replacement of, for example, a Landsat-7 band 1 (ETM1) image (whose spectral domain is 0.45–0.52 μm) with an (down-sampled) image of the same surface area acquired by an IKONOS band 1 sensor (whose spectral domain is 0.445–0.516 μm) should leave the LSRC output map unaffected (refer to Table II).

Due to the ISRC loss in spectral resolution with respect to LSRC's, the following is true.

1) The number of spectral types detected by ISRC in comparison with LSRC reduces from 46 to 25, which is approximately equal to a 45% loss [22]. The ISRC set of parent spectral categories becomes the following: 1) *'either snow or ice or cloud or bright bare soil or built up'* (SN_CL_BBB); 2) *'either water or shadow'* (WASH); 3) *'vegetation'* (V); 4) *'either bare soil or built up'* (BB); and 5) *'outliers'*.
2) Based on existing literature, ISRC is expected to be less effective than LSRC in the separation of vegetation from bare soil due to the loss of the Landsat MIR1 and MIR2 channels (refer to Table II) [22] and in the separation of snow from clouds and light-toned bare soils due to the loss of the MIR and TIR channels (refer to Table II) [22], [29]. In [25], to remove classification differences due to changes in radiometric calibration and SR of input sensors, ISRC and LSRC were compared in the classification of, respectively, an artificial (synthetic) four-band IKONOS-like image and its real seven-band Landsat-7 ETM+ master image. In that paper, the ISRC loss in classification accuracy in comparison with LSRC's was estimated at about 10%. In the same paper, the ISRC loss of accuracy was about 10% in the generation of a dichotomous vegetation/nonvegetation classification map. However, if spectral categories featuring sparse vegetation were not considered in this comparison, the ISRC capability of detecting vegetation was 99% consistent with that of LSRC, which is a result far better than what is theoretically expected [22], [25].

According to [25], its operational properties, namely, mapping accuracy, computational efficiency, degree of automation, and robustness to changes in the input data set, make the ISRC eligible for use as the preliminary pixel-based classification first stage in an operational two-stage stratified hierarchical RS-IUS instantiation that is suitable for mapping spaceborne VHR optical imagery [22], [25].

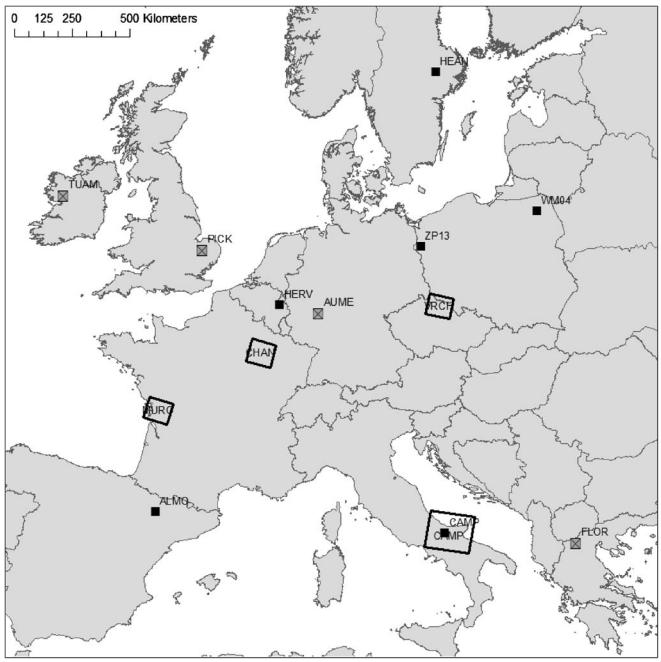


Fig. 2. Geographic footprints of the testing images identified by names of the testing sites and symbols, namely, 1) black square for (six) QB2 imagery, 2) gray square with black cross for (four) IK2 imagery, and 3) rectangle outline for (four) Landsat-5 TM (LTM) images. At the camp site, one QB2, one IK2, and one LTM images are available (refer to Table III).

III. Study Areas and Testing Image Set

The primary objective of this paper is the assessment of the ISRC accuracy and robustness to changes in the input data set acquired across time, space, and sensors at continental scale, with special emphasis on the detection of vegetation types (refer to Section I). According to the experimental session validity criteria proposed by Prechelt [55], at least two real and standard/appropriate (e.g., synthetic) data sets must be adopted to demonstrate the potential utility of an algorithm.

To reach the aforementioned paper's objective in agreement with the Prechelt experimental session validity criteria [55], a testing image set was collected according to the following constraints. 1) It is representative of the whole range of soil types encountered across the EU agricultural landscapes according to the European Soil Database v2.0 defined in agreement with the FAO/UNESCO classification system [56], [57]. These selected soil types exclude, for example, arid regions, mountainous areas, and urban areas. 2) It comprises both synthesized and real-world RS imagery. The former data set allows a comparison of the ISRC against the LSRC as a function of the change in spectral resolution, while radiometric calibration and SR do not change in the comparison. 3) It features an observation timing that is limited to the spring season, ranging from early April to late June, to deal with RS images dominated by the presence of green vegetation. 4) It avoids extreme off-nadir solar geometry. 5) Whenever possible, it adopts PAN-sharpening

TABLE IV
TEST SITES DEPICTED IN VHR IMAGES AND THEIR SOIL PROPERTIES TAKEN FROM THE EUROPEAN SOIL DATABASE v2.0 IN AGREEMENT WITH THE FAO/UNESCO CLASSIFICATION SYSTEM [56], [57]

Site identifier	Site name	Main soil types	Surface texture	Parent material
1	ZP13	Cambisol / podzol / arenosol	coarse + medium	quaternary clay + glacial sand
2	WM04	Cambisol / podzol / arenosol	coarse + fine	boulder clay + glacial sand
3	HEAN	Cambisol	medium	lake deposits
4	HERV	Luvisol / cambisol	medium fine	loess + stony loam
5	CAMP	Cambisol / vertisol	medium fine	sandstone + tertiary clay
6	ALMO	Cambisol / lithosol	medium + fine	marl
7	FLOR	Lithosol / fluvisol / luvisol	medium	limestone
8	PICK	Arenosol	coarse	cover sand
9	TUAM	Luvisol	medium	detrital limestone
10	AUME	Cambisol / luvisol	medium	metamorphic rocks + residual loam

to make SR finer, and thus, it reduces the presence of "mixed" pixels.

To satisfy the aforementioned constraints, 14 testing images were selected and preprocessed as described in the following. They comprise four radiometrically calibrated synthesized VHR-like images and ten radiometrically calibrated multisource real-world VHR images described in Table III. Their geographic footprints are shown in Fig. 2.

The four synthetic four-band IKONOS-like images shown in Fig. 2 represent the Mediterranean, Atlantic, and Central European condition of agricultural landscapes. They were generated from four "master" seven-band 30-m-resolution Landsat-5 TM scenes, radiometrically calibrated into TOARF values, whose channels TM1–TM4 provide the required IKONOS-like B, G, R, and NIR channels (refer to Table II).

The ten VHR images of agricultural landscapes shown in Fig. 2 comprise six 2.44-m-resolution QB2 images and four 4-m-resolution IK2 images. According to the European Soil Database v2.0 defined in agreement with the FAO/UNESCO classification system [56], [57], these VHR images cover the entire range of very weak, weak, moderate, and strong soil developments found across the EU, they have a surface mineral texture ranging from fine to coarse, and they exhibit a range of parent material mineralogical compositions that are typical of EU agricultural landscapes. These comments are summarized in Table IV.

Preprocessing of the selected QB2 and IK2 images was as follows. 1) Every testing image, provided with its radiometric calibration metadata file, was radiometrically calibrated into TOA radiance and, next, TOARF values [42]. 2) Four subsets, 500×500 pixels in size, were extracted from each VHR image to capture the diversity of that scene, contained a substantial portion of both vegetation and nonvegetation cover types, and included a significant portion of nonagricultural cover types in the vegetation part. 3) Each subset was independently PAN-sharpened at, respectively, 0.61-m resolution for QB2 and 1-m resolution for IKONOS imagery to generate an upscaled subimage of 2000×2000 pixels in size by means of the Gram–Schmidt PAN-sharpening algorithm implemented in the Environment for Visualizing Images commercial software toolbox licensed by ITT Industries, Inc. [58]. 4) Although the Gram–Schmidt algorithm supposedly preserves the spectral integrity of the original MS data, each PAN-sharpened image subset was compared with its original subimage. If a significant (linear) spectral distortion was found in the comparison, then a dark-object subtraction technique, typically used to remove the additive atmospheric scattering (haze) effects [59], was applied to the PAN-sharpened image subset. Only eight of 40 PAN-sharpened image subsets required a dark-object subtraction. Examples of a 0.61-m-resolution PAN-sharpened QB2 image and its ISRC map are shown in Figs. 3–5.

For each of the four subimages, 2000×2000 pixels in size, extracted from each of the ten testing VHR images, 300 reference pixels were randomly selected, i.e., 1200 reference samples were selected overall for each VHR testing image. Visually inspected by an expert photointerpreter, these reference pixels were mapped into the following three classes: vegetation, nonvegetation, and mixed pixels (see Table V, where the presence of mixed pixels ranges from 1% to 21% at subset level and from 2 to 12% at image level). High shares of mixed pixels are, to a large extent, due to emerging annual crops in early spring imagery.

IV. EXPERIMENTAL RESULTS

Rather mild algorithm benchmarking rules found in computer science literature are the following [55]: At least two real and standard/appropriate data sets must be adopted to demonstrate the potential utility of an algorithm (refer to Section III earlier), the proposed algorithm must be compared against at least one existing technique, and at least one fifth of the total paper length should be devoted to evaluation.

According to [22], LSRC (respectively, ISRC) maps each pixel of a radiometrically calibrated seven-band Landsat-like (respectively, four-band IKONOS-like) image into a set of 46 (respectively, 25) spectral categories (refer to Section II-B). Among these spectral categories, 14 (respectively, 13) spectral categories belong to "green vegetation" types, i.e., they are unambiguously identified as containing photosynthetically active vegetation.

From a theoretical point of view, due to the loss of the MIR channels, the downscaled four-band ISRC is less reliable than

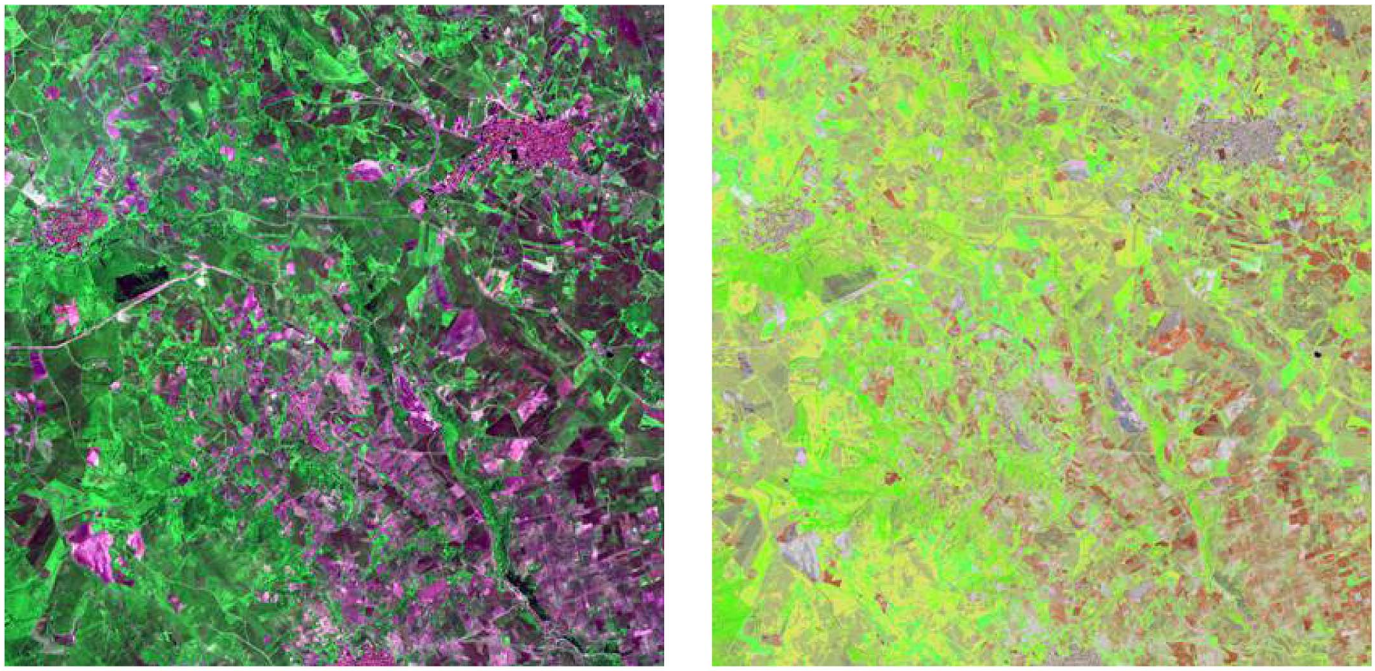


Fig. 3. (a) QuickBird-2 image of Campania, Italy (acquisition date: June 13, 2004, 09:58 GMT), radiometrically calibrated into TOARF values, PAN-sharpened from 2.44 to 0.61 m resolution and depicted in false colors (R: band 3, G: band 4, B: band 1). (b) ISRC output map depicted in pseudocolors and generated from the MS image shown in (a). The adopted pseudocolors are the following: green tones for vegetation and rangeland, brown and gray color shades for barren land and built-up areas, blue tones for water types, white tones for clouds and snow, and red tones for unknowns.

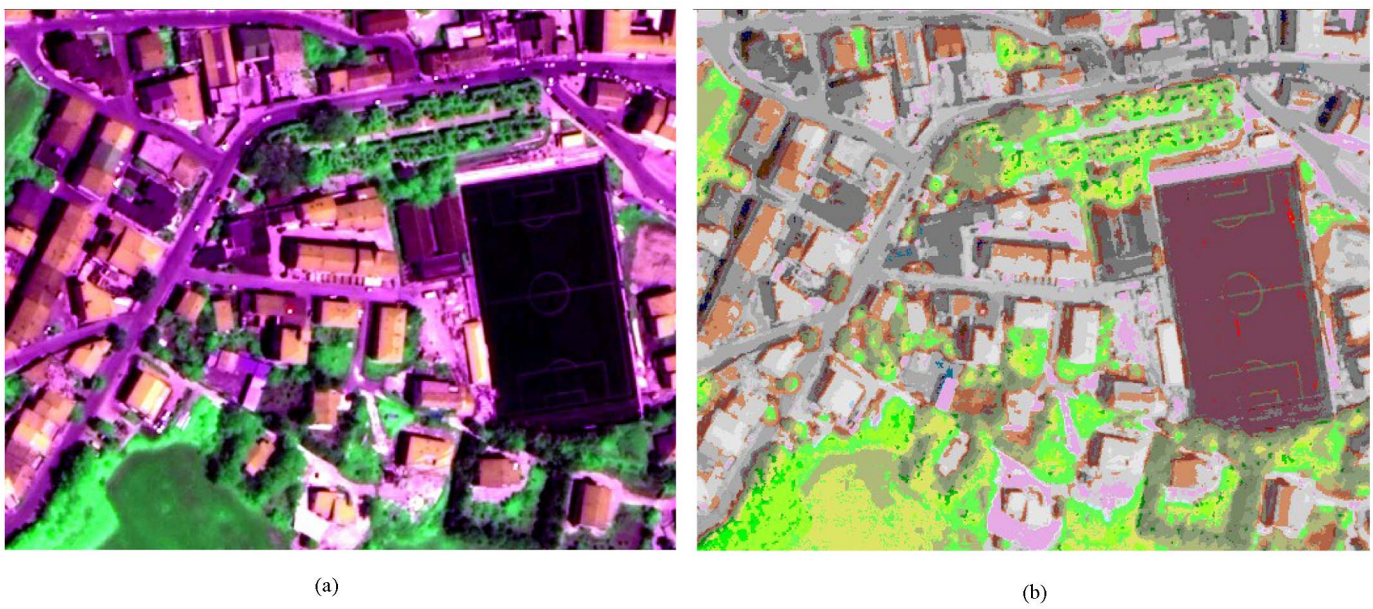


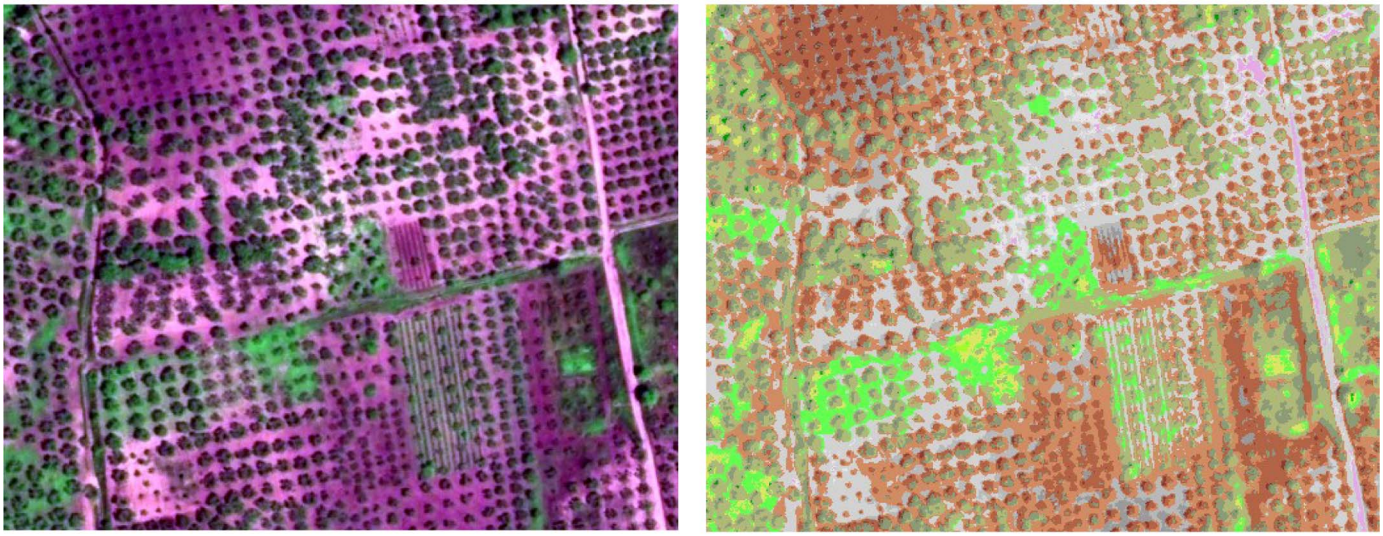
Fig. 4. (a) Zoomed image extracted from Fig. 3(a), radiometrically calibrated into TOARF values, PAN-sharpened from 2.44 to 0.61 m resolution and depicted in false colors (R: band 3, G: band 4, B: band 1). (b) Zoomed ISRC output map depicted in pseudocolors, extracted from Fig. 3(b) and overlapping with Fig. 4(a). Adopted pseudocolors are the same as those in Fig. 3(b).

the original seven-band LSRC in the detection of vegetation (see Section II-B) [22]. Fortunately, early experiments revealed an ISRC vegetation mapping accuracy superior to what is theoretically expected [25].

In this experimental section, the ISRC capability of mapping vegetation types is quantitatively assessed at the European continental scale to check the validity of preliminary conclusions gathered at local spatial extensions and reported in [25].

To reach this experimental objective and to fulfill the algorithm assessment criteria listed earlier, the following four experiments were conducted.

1) The first experiment quantitatively assesses the expected loss in spectral discrimination capability affecting the downscaled ISRC system in comparison with the original LSRC due to the inferior spectral resolution of the former irrespective of intersensor differences in radiometric



(a) (b)

Fig. 5. (a) Zoomed image extracted from Fig. 3(a), radiometrically calibrated into TOARF values, PAN-sharpened from 2.44 to 0.61m resolution and depicted in false colors (R: band 3, G: band 4, B: band 1). (b) Zoomed ISRC output map depicted in pseudocolors, extracted from Fig. 3(b) and overlapping with Fig. 5(a). Adopted pseudocolors are the same as those in Fig. 3(b).

calibration and SR (see Section III). To reach this goal, in line with [25], the ISRC is compared against the LSRC in the automatic generation of a binary vegetation mask extracted from, respectively, four synthesized IKONOS-like images (refer to Section III) and their four master Landsat-5 TM images (refer to Table III and Fig. 2).

2) The classification accuracy of the binary vegetation mask automatically generated by ISRC is estimated, employing, as input, 40 subsets of ten real-world PAN-sharpened QB2 and IK2 images (refer to Tables III and IV and Fig. 2). Each subset is provided with a reference data set of 300 samples (refer to Section III and Table V).
3) A traditional empirical scene-by-scene normalized difference vegetation index (NDVI) thresholding approach is compared against the fully automated ISRC in the detection of a dichotomous vegetation/nonvegetation mask across the 40 PAN-sharpened VHR image subsets that are 2000 $\times$ 2000 pixels in size (refer to Section III).
4) In an automatic two-stage stratified hierarchical RS-IUS instantiation employing ISRC as its preliminary classification first stage (refer to Section II), the second-stage stratified texture properties are extracted from the symbolic discrete ISRC map domain in place of the ordinary subsymbolic continuous varying MS image feature domain to separate low-texture *annual cropland or herbaceous rangeland* (identified as class AC/HR) from high-texture *forest or woodland* (identified as class F/W).

In this experimental session, the classification accuracy was assessed by means of the well-known classification overall accuracy (OA) probability that is provided with its error tolerance $p_{\mathrm{OA}} \in [0,1] \pm \delta$, with $0 < \delta < p_{\mathrm{OA}} \leq 1$ [75], in agreement with the QA4EO guidelines [49].

TABLE V
TESTING VHR IMAGES. CARDINALITY OF RANDOMLY SELECTED GROUND TRUTH PIXELS BELONGING TO CLASS VEGETATION, NONVEGETATION, OR MIXED PIXELS

VHR site name	Green vegetation dominated pixels	Mixed pixels	Pixels without significant green vegetation	Total
ZP13	494	147	559	1200
WM04	824	48	327	1200
HEAN	587	100	513	1200
HERV	837	42	321	1200
CAMP	530	105	565	1200
ALMO	800	88	312	1200
FLOR	369	129	702	1200
PICK	820	74	306	1200
TUAM	804	34	362	1200
AUME	913	26	261	1200

In line with the USGS classification system constraints, in this paper, the target $p_{\mathrm{OA}} \in [0,1]$ is fixed at 0.85, whereas the per-class classification accuracies should be about equal [52] and never below 70% [60].

Any p_{OA} estimate is a random variable (sample statistic) with a confidence interval (error tolerance) $\pm\delta$ associated with it. In other words, p_{OA} is a function of the specific testing data set used for its estimation [61]. It is easy to prove that, for a given $p_{\mathrm{OA}} \pm \delta$ value pair, a recommended cardinality of the testing data set M_{test} can be estimated. In practice, based on the assumption that reference samples are independent and identically distributed (i.i.d.; this assumption is typically violated in image classification problems due to spatial autocorrelation between neighboring reference pixels), for a given reference testing sample set size M_{test} and an estimated classification accuracy probability p_{OA} at a desired confidence level (e.g., if confidence level $= 95\%$, then the critical value is 1.96), the half

TABLE VI
COMPARISON OF ISRC WITH LSRC IN A DICHOTOMOUS VEGETATION/NONVEGETATION CLASSIFICATION PROBLEM. REFERENCE CLASSIFICATION MAPS ARE GENERATED BY LSRC FROM FOUR RADIOMETRICALLY CALIBRATED LANDSAT-5 TM MASTER IMAGES. TESTING CLASSIFICATION MAPS ARE GENERATED BY ISRC FROM FOUR ARTIFICIAL 4-BAND IKONOS-LIKE SLAVE IMAGES SYNTHESIZED FROM THE 7-BAND LANDSAT-5 TM MASTER IMAGES. OVERALL CLASSIFICATION ACCURACY PROBABILITY COMPUTED IMAGE-WIDE $= p_{\mathrm{OA}}$ (%). ERROR TOLERANCE AT A CONFIDENCE LEVEL OF $95\% = \pm\delta(\%) = (1)$.

Landsat site name	Rows	Columns	Veg./ Non-veg. p_{OA}	$\pm\delta$ = (1)	Veg. Omission	Veg. Commission
CAMP	7617	7023	82.75	0.010%	17.03	0.22
CHAN	7620	7024	97.34	0.004%	1.45	1.21
MURO	7828	7271	92.46	0.007%	6.87	0.67
VRCH	7888	7289	76.18	0.011%	17.89	5.93

TABLE VII
COMPARISON OF ISRC WITH LSRC IN A DICHOTOMOUS VEGETATION/NONVEGEATATION CLASSIFICATION PROBLEM. SELECTION OF SPECTRAL CATEGORIES DETECTED BY LSRC (EMPLOYED AS A REFERENCE) AND ISRC THAT OVERLAP WITH A LARGE PORTION (SHARE) OF THE ISRC VEGETATION OMISSION ERRORS REPORTED IN TABLE VI

Landsat site name	LSRC spectral types of interest (% = Share of ISRC vegetation omission errors overlapping with the selected LSRC spectral types)	ISRC spectral types of interest (% = Share of ISRC vegetation omission errors overlapping with the selected ISRC spectral types)
CAMP	WR, WEDR (97.0 %)	SBBNF, ABBNF (99.8 %)
CHAN	WR, WEDR (85.5%)	SBBNF, ABBNF (99.3 %)
MURO	WR, WEDR (88.1 %)	SBBNF, ABBNF (98.7 %)
VRCH	WVLNIR, WR, WEDR (84.1 %)	SBBNF, ABBNF (94.4 %)

width δ of an error tolerance that is equal to $\pm\delta$ becomes [43], [62], [63]

$$\delta = \sqrt{\frac{(1.96)^2 \cdot p_{\mathrm{OA}} \cdot (1 - p_{\mathrm{OA}})}{M_{\mathrm{test}}}}. \tag{1}$$

According to (1), in a dichotomous vegetation/nonvegetation classification problem, if the target $p_{\mathrm{OA}} = 0.85$ (see this text earlier) with $M_{\mathrm{test}} = 1200$ (refer to Section III) and if confidence level $= 95\%$, then $\delta = (1) = 2\%$.

For each cth class simultaneously involved in the classification process, with $c = 1, \ldots, C$, it is possible to prove that (refer to [43, p. 294])

$$\delta_c = \sqrt{\frac{\chi^2_{(1,1-\alpha/C)} \cdot p_{\mathrm{OA},c} \cdot (1 - p_{\mathrm{OA},c})}{m_c}}, \qquad c = 1, \ldots, C \tag{2}$$

where α is the desired level of significance, i.e., the risk that the actual error is larger than δ_c (e.g., $\alpha = 0.03$), $1 - \alpha/C$ is the level of confidence, and $\chi^2_{(1,1-\alpha/C)}$ is the upper $(1 - (\alpha/C)) * 100$th percentile of the chi-square distribution with one degree of freedom. For example, if $\alpha = 0.03$ and $C = 3$, then the level of confidence is $(1 - 0.03/3) = 0.99$ and $\chi^2_{(1,0.99)} = 6.63$. In this case, if $p_{\mathrm{OA},c} = 0.70$ (see this text earlier), $m_c = 300$ (refer to Section III), and confidence level $= 95\%$, then $\delta_c = (2) = 6.8\%$.

A. Quantitative Assessment of ISRC Accuracy

Experiments 1)–4) that are described in the introduction of Section IV are discussed in the following.

1) Synthesized IKONOS-Like Images at Regional Scale: To remove the differences in sensor calibration and SR, ISRC is compared against LSRC in the automatic generation of a binary vegetation mask from, respectively, four synthesized IKONOS-like slave images and their four Landsat-5 TM master images (refer to Table III).

In these four experiments, column 4 of Table VI shows that, when compared to LSRC, ISRC provides a binary vegetation mask whose p_{OA} value ranges from 76% to 97%. Columns 6 and 7 of Table VI reveal that the vegetation mapping difference between the reference LSRC and the testing ISRC is mainly due to omission errors (type I errors [13]). Table VII investigates the degree of semantic shift of the ISRC vegetation omission errors reported in Table VI from the LSRC reference. In particular, Table VII shows that the ISRC vegetation omission errors reported in Table VI are mainly due to pixels whose reference LSRC spectral types feature a low leaf area index (LAI), such as spectral categories *'weak rangeland'* (WR) and *'either wetland or dark rangeland'* (WEDR). These low LAI value-pixels are erroneously mapped by ISRC onto nonvegetated spectral types such as "non flat" (NF) bare soil or built-up (BB) [e.g., *'strong bare soil or built-up non flat'* (SBBNF) and *'average bare soil or built-up non flat'* (ABBNF)]. To summarize, Tables VI and VII together show that, when the seven-band LSRC is adopted as a reference, the downscaled four-band ISRC mislabels either "mixed pixels" with a sparse vegetation cover or pixels featuring a full vegetation cover but a low chlorophyll content. It is noteworthy that, if pixels labeled by the LSRC as featuring sparse vegetation cover or a low LAI value (e.g., belonging to spectral types WR and WEDR) are ignored in this comparison, then, according to Tables VI and VII, the degree of match between the reference LSRC and the testing ISRC binary vegetation maps exceeds 98.5%, which is perfectly in line with the 99% degree of match reported in [25] in a similar experiment conducted at local scale (refer to Section II-B).

Column 7 in Table VI shows that the ISRC binary vegetation mask commission errors are always below 6% and that, in general, they appear negligible in terms of occurrence. The semantic shift associated with the ISRC vegetation commission

TABLE VIII
COMPARISON OF ISRC WITH LSRC IN A DICHOTOMOUS VEGETATION/NONVEGETATION CLASSIFICATION PROBLEM. SELECTION OF SPECTRAL CATEGORIES DETECTED BY LSRC (EMPLOYED AS A REFERENCE) AND ISRC THAT OVERLAP WITH A LARGE PORTION (SHARE) OF THE ISRC VEGETATION COMMISSION ERRORS REPORTED IN TABLE VI

Landsat site name	LSRC spectral types of interest (% = Share of ISRC vegetation commission errors overlapping with the selected LSRC spectral types)	ISRC spectral types of interest (% = Share of ISRC vegetation commission errors overlapping with the selected ISRC spectral types)
CHAN	TNCLV, UN (71.9%)	AVVHNIR, AHR, PB (58.6%)
MURO	TNCLV, UN (44.8%)	ASRMNIR, PB (39.3%)
VRCH	TNCLV (87.7%)	ASRHNIR, ASRMNIR, PB (30.0%)

errors reported in Table VI is investigated in Table VIII. It reveals that, first, mixed pixels labeled as '*outliers*' (unknown, UN) by the reference LSRC can be mapped by ISRC onto vegetation spectral types. The frequency of these pixels in the testing images appears extremely low; thus, their impact on the LSRC and ISRC vegetation mapping difference is negligible. Second, a large portion of the reference LSRC spectral category *'thin cloud over vegetation'* (TNCLV) is identified by ISRC as vegetation. Although this is a semantic shift between ISRC and LSRC mappings, it cannot be considered as a vegetation overestimation of ISRC with respect to LSRC.

To summarize, Tables VI–VIII show first that, in four synthesized VHR images at regional scale, ISRC maintains the LSRC capability of detecting vegetation types (whose LAI ranges from low to high) from 76% to 97%, which rises to about 99% if the pixels with sparse vegetation cover (low LAI) are not considered in the comparison. Second, in comparison with LSRC, ISRC is affected by an omission error (ranging from 1% to 17%) when the canopy is sparse or when the chlorophyll content is low, while its commission error tends to be always small ($< 6\%$). These conclusions are perfectly in line with the theoretical expectations. Although forgotten in RS common practice, it is well known in existing literature that the NDVI (equivalent to a first derivative of the spectral signature between the visible and NIR portions of the electromagnetic spectrum), which is also, but not exclusively, employed by the ISRC decision rules, is ineffective when the vegetation canopy is too dense or too sparse [64, p. 250]. In addition, the aforementioned conclusions holding at regional scale generalize preliminary observations collected at local scale in [25], where the ISRC capability of detecting vegetation was 99%, which is consistent with that of the LSRC when the spectral categories featuring sparse vegetation were omitted from the classification comparison (refer to Section II-B).

In RS common practice, this paper demonstrates that the binary vegetation masks provided by LSRC and ISRC feature different semantic meanings. While the LSRC vegetation mask separates pixels containing no vegetation at all from pixels containing any degree of green vegetation ranging from low to high, the ISRC vegetation mask separates pixels containing either a large or medium amount of green vegetation from pixels containing either no or little chlorophyllous biomass.

2) PAN-Sharpened IK2 and QB2 Images at Local Scale: The classification accuracy of the binary vegetation mask generated by the fully automated ISRC was estimated across the 40 PAN-sharpened VHR image subsets, 2000 $\times$ 2000 pixels in size (refer to Tables III and IV and Fig. 2). For each image subset, the reference data set comprises 300 randomly selected pixels assigned to class vegetation, nonvegetation, or mixed pixels (refer to Section III and Table V).

At a visual assessment, the overall quality of the preliminary classification maps generated as output by ISRC from the four PAN-sharpened 1-m-resolution IK2 and the six PAN-sharpened 0.62-m-resolution QB2 images selected for testing appears very satisfactory (see Figs. 3–5). It is noteworthy that, unlike traditional pixel-based classifier such as ML and neural network classifiers [15], the automatic pixel-based model-driven ISRC appears affected by no salt-and-pepper classification noise effect. This is due to the ISRC capacity of modeling within-spectral category variance [22], [25].

The accuracy of a classification map is expected to increase when the SR becomes finer, i.e., when the number of mixed pixels decreases [65]. This means that the mapping accuracy of the ISRC binary vegetation masks generated in this section at an SR below or equal to 1 m is expected to be not inferior to that estimated in Section IV-A1, dealing with 30-m-resolution ISRC binary vegetation masks.

Table IX provides the ISRC binary vegetation mask p_{OA} estimate per site and the average binary vegetation/nonvegetation p_{OA} value computed across the 40 ISRC binary vegetation masks (i.e., without considering reference mixed pixels), which is equal to 98.21%, with an average error tolerance equal to $\pm 0.79\%$. This outcome is in line with previous results gathered with synthetic IKONOS-like images in Section IV-A1. Columns 4 and 5 of Table IX show that, at subset level, for both vegetation and nonvegetation reference data sets, the share (%) values of correctly labeled pixels are in all cases, but one, equal to or above 88%. The only exception, which is a nonvegetation reference data set whose correctly mapped percentage value is equal to 70.6%, occurs with one subset in site HERV. A visual inspection of this image subset revealed that random nonvegetation reference samples were located on the edges of small cirrus clouds overlying a uniformly vegetated area. These pixels were labeled as nonvegetation (namely, cloud) by an expert photointerpreter, but the strong signal of the underlying vegetation is detected through thin clouds by ISRC, which assigns these pixels with spectral category label *'vegetation'* (also refer to a similar circumstance that occurred in Section IV-A1). Although this is a semantic shift from reference samples, it cannot be considered a vegetation overestimation error by ISRC.

Columns 6 and 7 of Table IX reveal that the population of the reference mixed pixels is mapped by ISRC into spectral categories vegetation or nonvegetation according to proportions that vary with site characteristics. This observation supports

TABLE IX
PAN-SHARPENED VHR TESTING IMAGES. ISRC BINARY VEGETATION MASK OA PROBABILITY PER SITE=p_{OA} (IN PERCENT). ERROR TOLERANCE AT A CONFIDENCE LEVEL OF $95\% = \pm\delta\ (\%) = (1)$. THE AVERAGE p_{OA} VALUE COMPUTED ACROSS THE 40 ISRC BINARY VEGETATION MASKS AT SUBSET LEVEL IS EQUAL TO $98.21\% \pm 0.79\%$

VHR site name	Two-class veg. / non-veg. p_{OA} value per site, ignoring reference mixed pixels. Mean value across sites = 98,21 (%)	±δ = Eq. (1). Mean value across sites = ± 0.79 (%)	Average and, between brackets, min and max share (%) of reference veg. pixels mapped correctly at subset level per site	Average and, between brackets, min and max share (%) of reference non-veg. pixels mapped correctly at subset level per site	Reference mixed pixels detected as vegetation (%)	Reference mixed pixels detected as non-vegetation (%)
ZP13	97,7	0,9	98,2 *(94.0-100.0)*	97,3 *(88.2-99.3)*	71,4	28,6
WM04	99,6	0,4	99,9 *(99.4-100.0)*	98,7 *(96.7-100.0)*	47,9	52,1
HEAN	98,8	0,6	99,7 *(98.9-100.0)*	97,8 *(96.1-100.0)*	80,0	20,0
HERV	97,2	1,0	99,6 *(99.1-100.0)*	90,5 *(70.6-98.7)*	81,0	19,0
CAMP	97,3	1,0	95,8 *(88.0-97.1)*	98,6 *(97.2-99.4)*	52,4	47,6
ALMO	97,0	1,0	96,3 *(93.2-99.1)*	99,0 *(95.6-100.0)*	61,4	38,6
FLOR	99,1	0,6	98,6 *(92.9-100.0)*	99,4 *(97.6-100.0)*	59,7	40,3
PICK	97,7	0,9	97,3 *(93.9-99.6)*	98,9 *(97.5-100.0)*	45,9	54,1
TUAM	98,5	1,0	99,5 *(99.4-99.6)*	95,9 *(94.3-100.0)*	91,2	8,8
AUME	99,2	0,5	100,0	96,5 *(91.1-98.6)*	53,8	46,2

the conclusion of Section IV-A1; according to it, the ISRC binary vegetation mask locates pixels featuring either a large or medium amount of green vegetation, while pixels containing either no or little chlorophyllous biomass (like mixed pixels generated from a mixture of vegetation with nonvegetation surface types) are masked out. It is noteworthy that, across columns 6 and 7 of Table IX, the only circumstance where ISRC mapping of mixed pixels is clearly unbalanced in favor of the vegetation class occurs at the TUAM site. A visual inspection of this testing image revealed that reference mixed pixels belong to a peat bog. Since a peat bog is a heavily organic soil where regenerating weeds may grow, it is intrinsically "fuzzy" to map. To summarize, this special case does not question the conclusion which states that ISRC appears unbiased when mapping mixed pixels into classes vegetation and nonvegetation.

3) Comparison of the ISRC Vegetation Mask With an NDVI-Based Vegetation Mask: Ordinary empirical scene-by-scene NDVI thresholding was compared against the fully automated ISRC in the detection of a dichotomous vegetation/nonvegetation mask across the 40 PAN-sharpened VHR image subsets. It is noteworthy that NDVI is computed from input MS imagery that is radiometrically calibrated into TOARF values, as recommended in [64, p. 271]. By a trial-and-error strategy, an empirical NDVI threshold was selected to be equal to 0.36 to maximize the average binary vegetation/nonvegetation p_{OA} value computed across the 40 binary vegetation masks. This maximum average p_{OA} value is equal to 97.68%, with an average error tolerance of $\pm 0.89\%$ at a confidence level that is equal to 95% (refer to Table X). It is slightly lower than the average p_{OA} value provided by the automatic ISRC in Table IX, which is equal to $98.21\% \pm 0.79\%$. The comparison of Table X with Table IX reveals that, at subset level, the binary vegetation mask p_{OA} values of the NDVI thresholding approach are slightly, but not significantly, less accurate than those generated by ISRC. This overall similarity is not entirely surprising since, to detect vegetation, both approaches rely heavily (in the case of ISRC) or exclusively (in the case of the NDVI thresholding technique) upon the same two-of-four spectral bands, namely, the visible R and the NIR channel, whereas visible B and visible G bands, which are highly correlated with band R, are neglected.

In spite of the aforementioned accuracy similarities between the ISRC and the NDVI thresholding approach, the latter has several operational disadvantages in comparison with ISRC. These operational disadvantages are summarized in the following.

1) The NDVI thresholding approach is applied on a scene-by-scene basis, i.e., it is empirically driven by data, while ISRC is automatic. Robustness of the NDVI thresholding approach to changes in the input data set remains unknown.
2) The NDVI thresholding approach is unable to cope with areas where the vegetation spectral reflectance model may not be ordinary, e.g., due to indirect illumination such as vegetation in shadow areas [66], [67].
3) The NDVI thresholding approach is affected by well-known commission errors, e.g., cirrus clouds and some types of bright bare soils may feature high NDVI values. In existing literature, it is well known that the NDVI alone, by measuring the contrast between channels NIR and R, is unable, *per se*, to guarantee a robust (i.e., reliable and image-independent) discrimination between vegetation and nonvegetation surface types [26], [64]. Unfortunately, this evidence is often forgotten in RS common practice [3].

An alternative NDVI thresholding approach would be to establish the NDVI threshold on the basis of the (empirical) square root relation between NDVI and the vegetation cover fraction, as proposed by [68] and adopted by [24] and [44]. According to the square root relation [68], an NDVI threshold value is selected to represent a (green) vegetation cover fraction of only 25%. In our experiments, when this alternative NDVI threshold was applied, vegetation mapping accuracy values

TABLE X
PAN-SHARPENED VHR TESTING IMAGES. THE SELECTED NDVI (THRESHOLD=0.36) MAXIMIZES THE AVERAGE VEGETATION p_{OA} VALUE COMPUTED ACROSS THE 40 BINARY VEGETATION MASKS AT SUBSET LEVEL, WHICH IS EQUAL TO 97.68%, WITH AN ERROR TOLERANCE AT A CONFIDENCE LEVEL OF $95\% = \pm\delta\ (\%) = (1) = \pm 0.89\%$

VHR site name	Two-class veg. / non-veg. pOA value per site, ignoring reference mixed pixels. Mean value across sites = 97,68 (%)	±δ = Eq. (1). Mean value across sites = ± 0,89 (%)	Share (%) of reference veg. pixels mapped correctly at subset level per site	Share (%) of reference non-veg. pixels mapped correctly at subset level per site	Reference mixed pixels detected as vegetation (%)	Reference mixed pixels detected as non-vegetation (%)
ZP13	98,0	0,9	97,2	98,8	66,0	34,0
WM04	98,7	0,6	98,7	99,4	41,7	58,3
HEAN	98,3	0,8	97,6	99,0	76,0	24,0
HERV	95,8	1,2	96,1	95,1	69,0	31,0
CAMP	97,2	1,0	95,3	98,9	48,6	51,4
ALMO	96,8	1,0	95,9	99,0	60,2	39,8
FLOR	97,5	0,9	93,8	99,6	45,0	55,0
PICK	96,8	1,0	96,1	98,9	37,8	62,2
TUAM	98,5	1,0	99,3	96,6	85,3	14,7
AUME	99,2	0,5	100,0	96,5	50,0	50,0

were less accurate than those shown in Table X (e.g., in the HEAN site, the vegetated pixel detection accuracy decreased from 99.7% by ISRC in Table IX to 97.6% by ordinary NDVI thresholding in Table X and to 90.8% by the second NDVI thresholding technique). In particular, the average p_{OA} value, which is equal to 98.21% in Table IX and 97.68% in Table X, decreases to 94.8% by the second NDVI thresholding technique. This result suggests that the (empirical) square root relation between the NDVI and the vegetation cover fraction does not hold in our experiments.

B. Second-Stage Stratified Context-Based Classification of Woodland and HR

A two-stage stratified hierarchical RS-IUS employing ISRC as its pixel-based preliminary classification first stage (see Section II) was instantiated to distinguish low-texture class AC/HR from high-texture class F/W in the symbolic, discrete, and finite preliminary classification map domain in place of the ordinary subsymbolic continuous varying image feature domain. In the RS literature, dozens of works separate vegetation types based on a nonstratified (unconditional and driven without knowledge) texture analysis in the MS image domain with varying degrees of success (e.g., refer to [69]). In contrast with these traditional texture detection approaches, Shackelford and Davis presented two-stage stratified hierarchical RS-IUS implementations whose second stage employs a stratified texture analyzer [19], [20]. However, neither the former nor the latter approaches are based on texture analysis in the (symbolic, discrete, and finite) preliminary classification map domain in place of the traditional (subsymbolic, continuous, color chromatic, or achromatic [16]) signal domain.

It is noteworthy that neither a formal definition nor a class-specific texture model based on invariant image properties holds for a textured land cover class such as "woodland," which encompasses many vegetation types featuring varying texture elements (textons), orientations, and periods [69]. To overcome this limitation, a multiscale texture-sensitive model-driven (top-down and learning by rule) second-stage F/W class-specific classification module can employ, as input, the symbolic and discrete preliminary ISRC map domain in place of the ordinary image feature space. In particular, the center of a multiscale local window is moved across the binary vegetation mask that is automatically detected by ISRC. For each window position centered on the binary vegetation mask, the distribution (first-order histogram) of the preliminary ISRC map labels, masked by the vegetation mask, is collected. This vegetation mask-conditional label distribution is expected to increase its complexity (number of bins with occurrence above zero, which is equivalent to a measure of entropy) monotonically with vegetation texture. In fact, the number of ISRC spectral categories that are capable of mapping the low-texture AC/HR class is expected to be inferior to the number of discrete spectral categories that are capable of mapping the high-texture F/W class.

To investigate this hypothesis, a stratified set of approximately nine randomly sampled reference pixels was selected in each of the four image subsets per testing site (refer to Section III) per class F/W and class AC/HR as follows. First, an image stratum is computed as the intersection between the binary vegetation mask generated from the VHR image subset at hand and the buffer zone image consisting of pixels located at least 15 m from the border of the image subset (to avoid boundary effects across a moving window centered on pixels belonging to the image stratum). Second, random samples are selected within the image stratum and at least 40 m from each other (to reduce reference sample autocorrelation). Finally, stratified random samples are manually labeled by an expert photointerpreter as belonging to either class AC/HR or F/W. Thus, approximately 350 ground truth pixels per class F/W (namely, 369 reference pixels) and class AC/HR (namely, 317 reference pixels) were selected overall across the ten VHR image testing sites (refer to Table XI).

Spectral category index first-order distributions, masked by the binary vegetation mask, were assessed at multiple scales

TABLE XI

LAND COVER CLASS-CONDITIONAL PRELIMINARY SPECTRAL CATEGORY INDEX DISTRIBUTIONS WITHIN 15-m-RADIUS NEIGHBORHOODS CENTERED ON THE BINARY VEGETATION MASK. TARGET LAND COVER CLASSES ARE THE FOLLOWING: 1) AC/HC AND 2) F/W. THE FREQUENCY OF A GIVEN SPECTRAL CATEGORY, IN RANGE [0, 1], IS EXPRESSED AS A FRACTION OF THE TOTAL SAMPLE SIZE. THE "MEAN OCCURRENCE PER LOCAL WINDOW" PRESENTS THE AVERAGE SHARE OF THE AREA OF THE 15-m-RADIUS CIRCLE OCCUPIED BY EACH SPECTRAL CATEGORY, AND "STD" PROVIDES THE STANDARD DEVIATION OF THIS ESTIMATE. NSC IS THE NUMBER OF SPECTRAL CATEGORIES. SDI IS THE SHANNON DIVERSITY INDEX

Window size = 15 m radius circle		ISRC spectral categories belonging to the vegetation mask													Texture statistics	
		SVVHNIR	SVHNIR	SVMNIR	SVLNIR	SVVLNIR	AVVHNIR	SHRWE	SHV_WEDR	ASRHNIR	ASRMNIR	ASRLNIR	ASRVLNIR	PB	NSC	SDI
AC/HR (reference pixel cardinality = N1 = 369)	Freq. in N1 local windows	0,37	0,09	0,00	0,00	0,00	0,67	0,00	0,00	0,57	0,18	0,04	0,00	0,48		
	Mean occurrence per local window	0,17	0,01	0,00	0,00	0,00	0,33	0,00	0,00	0,20	0,08	0,00	0,00	0,19	2,48	0,39
	Std	0,34	0,06	0,00	0,00	0,00	0,38	0,00	0,00	0,32	0,25	0,05	0,00	0,32	1,03	0,33
F/W (reference pixel cardinality = N2 = 317)	Freq. in N2 local windows	0,26	0,34	0,42	0,37	0,14	0,34	0,09	0,27	0,73	0,91	0,81	0,59	0,53		
	Mean occurrence per local window	0,03	0,03	0,01	0,01	0,00	0,07	0,00	0,02	0,13	0,23	0,26	0,09	0,07	7,57	1,18
	Std	0,10	0,06	0,02	0,02	0,00	0,17	0,01	0,06	0,18	0,20	0,26	0,14	0,16	2,90	0,34

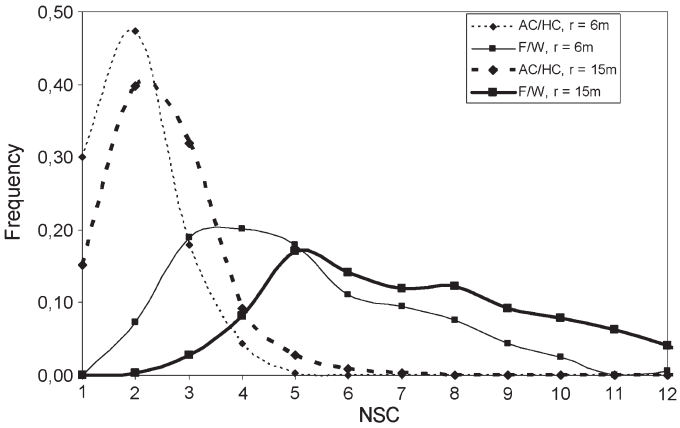


Fig. 6. Distribution of the NSC detected within 6- and 15-m-radius neighborhood overlapping with vegetation spectral categories detected by IRSC per class: 1) AC/HR and 2) F/W.

to provide an effective decomposition of the scene [16], [70]–[72], and according to the *uncertainty principle* (refer to Section II-A), they guarantee the best tradeoff between accurate boundary localization and the size (activation domain and receptive field) of the local window used to compute contextual features [13], [73]. By a trial-and-error strategy, it was observed that a neighborhood radius beyond 15 m did not enhance the collected texture information. Thus, two near dyadic circles featuring a radius of, respectively, 6 and 15 m were selected (see Fig. 6). The small circle with a 6-m radius provides the minimum representative sample size, and it was considered to be equivalent to the texture period of the low-texture class AC/HR, whereas a neighborhood with a 15-m radius was selected as a representative of the texture period of the high-texture class F/W.

Table XI presents the average frequency values computed across the ten testing sites of the 13 ISRC vegetation spectral categories falling within a moving window with a 15-m radius and centered on every reference sample belonging to classes AC/HR and F/W. It is noteworthy that 6 of the 13 vegetation spectral categories never occur in class AC/HR. This observation holds true also for the 6-m-radius neighborhood statistics. For more than half of the F/W class observations, 5 of the 13 vegetation spectral indices occur, with their individual average occurrence per local window ranging from 7% to 26%. One of these five indices (namely, ASRVLNIR) never occurs in class AC/HR. These considerations account for the inferior texture content of class AC/HR with respect to class F/W, in line with theoretical expectations (refer to this text earlier).

In the last two columns on the right of Table XI, vegetation-masked first-order label distributions are parameterized by alternative "diversity" (entropy) indicators, such as the Shannon diversity index (SDI) [74] and the simple number of spectral categories (NSC), whose cardinality is above zero. The latter performs better than the former in terms of sensitivity to the target texture phenomenon.

Fig. 6 shows the frequency distributions of NSC for the two window radii and the two target land cover classes AC/HR and F/W. Moving from 6- to 15-m radius, the maximum of the F/W class-conditional NSC distribution increases from NSC $=$ 4 to 5, while the maximum for class AC/HR remains equal to NSC $= 2$. At a 15-m radius, the following is true: 1) If NSC ≥ 4, then 3% of the F/W samples are omitted; 2) if NSC > 4, then only 3% of the AC/HR samples are detected; and 3) if NSC $= 4$, less than 9% of the F/W and AC/HR samples are selected. In the latter case, it is observed that vegetation spectral labels SVVHNIR or SVHNIR (equivalent to large LAI values)

appear in 23 of the 35 observations for class AC/HR and 3 of the 26 observations for class F/W.

Starting from these observations, Table XII presents the pseudocode of a second-stage stratified context-sensitive hierarchical rule-based classifier to be employed in series with the first-stage ISRC. This two-stage classifier separates vegetation land cover classes F/W and AC/HR based on texture properties computed in a symbolic discrete preliminary spectral classification map domain rather than in an ordinary continuous varying texture feature space. In ten VHR images acquired at European scale, this two-stage stratified hierarchical RS-IUS is accurate (featuring per class omission and commission errors $\leq 6\%$) and computationally efficient (near real time).

V. Summary and Conclusion

In recent years, the number of VHR spaceborne platforms and the demand for VHR commercial satellite imagery have continued to increase in terms of both quantity and quality of data. This has driven the need for automating quantitative analysis of VHR imagery. Unfortunately, existing two-stage segment-based RS-IUSs, currently considered as the state of the art in spaceborne VHR image processing software toolboxes, score poorly in terms of operational performance.

Downscaled from an original seven-band LSRC system, the automatic near-real-time four-band ISRC has recently been presented as the pixel-based preliminary classification first stage of a two-stage stratified hierarchical RS-IUS instantiation that is suitable for mapping VHR imagery and that is an alternative to traditional two-stage segment-based RS-IUSs.

The consequences of the ISRC loss in spectral resolution with respect to LSRC's are twofold: 1) The number of spectral types detected by ISRC in comparison with LSRC reduces from 46 to 25, which is approximately equal to a 45% loss [22], and 2) the following is expected from the ISRC: a) it features low sensitivity in the spectral discrimination of nonvegetation pixels (featuring LAI values that are approximately equal to zero) from pixels featuring a mixed (sparse) and/or weak green vegetation presence (featuring low LAI values), and b) it can be affected by a spectral overlap between classes cloud, snow, and light-toned bare soil. In [25], a preliminary quantitative assessment of the ISRC vegetation mapping capability at local scale showed a degree of consistency with the LSRC output map, which is far superior to what is theoretically expected.

This paper provides a so-called stage 1 validation of the ISRC vegetation/nonvegetation mapping capability in terms of time sampling, limited to the spring season, and a so-called stage 2 validation of the ISRC vegetation/nonvegetation map products in terms of space sampling, extended to the European continent. At regional scale, the selected testing data set comprises four artificial IKONOS-like images synthesized from Landsat-5 TM images and, at local scale, ten multisource VHR IK2 and QB2 images of agricultural landscapes selected across the EU in the spring season (from early April to late June).

ISRC is compared against LSRC in the classification of the four selected synthetic VHR images. This comparison reveals that the degree of consistency of the ISRC vegetation map with LSRC's ranges from 76% to 97%, mainly due to omission errors caused by mixed pixels corresponding to sparse and/or weak green vegetation (featuring low LAI values). Whereas these pixels are mapped onto the vegetation mask by LSRC, they are mapped onto the nonvegetation mask by ISRC. This means that the dichotomous vegetation/nonvegetation map detected by ISRC separates either strong or average vegetation pixels from either weak or no vegetation pixels, unlike LSRC which separates pixels featuring a degree of vegetation ranging from low to high from nonvegetation pixels. These results are perfectly in line with those collected at local scale in [25].

In the classification of ten real-world PAN-sharpened VHR images, the average ISRC dichotomous vegetation/nonvegetation detection accuracy computed across sensors and testing sites without considering mixed pixels was estimated at 98.21%, with an average error tolerance of $\pm 0.79\%$. This result is in line with the results collected with the synthetic image set. It is also measurably superior compared to that generated by an ordinary scene-specific NDVI thresholding technique in terms of degree of automation, mapping accuracy, and robustness to changes in the input image acquired across sensor and space.

To summarize, the aforementioned experimental results demonstrate that, although ISRC is affected by theoretical limitations in the spectral discrimination of vegetation land cover classes with respect to the seven-band LSRC due to the loss of the MIR and TIR channels, ISRC can be considered effective and robust to changes in the input VHR image of agricultural landscapes selected across the EU.

Another experiment highlights the degree of novelty of the proposed two-stage stratified hierarchical RS-IUS architecture, employing ISRC as its pixel-based preliminary classification first stage that is suitable for mapping VHR imagery. In series with ISRC, a second-stage land cover class-specific classification module is implemented as a context-sensitive decision rule employing, as input, the symbolic discrete and finite preliminary ISRC map in place of a traditional, subsymbolic, and continuous varying texture feature space to discriminate between textured land cover classes, namely, low-texture AC/HR from high-texture F/W. In this experiment, the implemented near real-time second-stage stratified context-sensitive rule-based classifier maintains omission and commission errors at about $3 \div 4\%$ of the reference data set.

The validated ISRC output products at the European continental scale may be relevant to the following: 1) the various levels and stages of the annual Control With RS program of the EU Common Agricultural Policy, which is still largely based on computer-assisted photo-interpretation techniques; 2) the automatic implementation of hierarchical RS data classification taxonomies, whose first decision level consists of a dichotomous vegetation/nonvegetation classification map, such as the CORINE, the USGS classification hierarchy, and the FAO LCCS; and 3) the implementation of operational GEOSS and GMES instantiations.

As a future development of this paper, a two-stage stratified hierarchical RS-IUS architecture employing ISRC as its pixel-based preliminary classification first stage will be provided with a battery of second-stage stratified class-specific classification

TABLE XII
PSEUDOCODE IN SERIES WITH ISRC TO SEPARATE CLASS F/W FROM CLASS AC/HR BASED ON STRATIFIED TEXTURE PROPERTIES IN THE PRELIMINARY ISRC MAP DOMAIN

```
//at this stage, i.e., at the output of the ISRC, pixel-based spectral properties have been exploited, exclusively.
IF (pixel p0 ∈ binary vegetation mask extracted from the preliminary ISRC map) THEN
{
    In a neighborhood centered on pixel p0 within a 15 m radius, compute the local histogram of spectral
    category indexes belonging to the preliminary spectral classification map masked by the vegetation mask,
    where NSC is the number of histogram bins with occurrence above 0.
    IF (NSC > 4) THEN
    {
        Land cover class is F/W.
        //estimated omission error = 9% (if NSC = 4, to be
        //recovered in part, see below) + 3% (if NSC < 4) = 12%.
        //estimated commission error (if NSC > 4, due to AC/HR) = 3%.
    }
    ELSE
    {
        IF (NSC < 4) THEN
        {
            Land cover class is AC/HR.
            //estimated omission error = 9% (if NSC = 4, to be
            //recovered in part, see below) + 3% (if NSC > 4) = 12%.
            //estimated commission error (if NSC < 4, due to F/W) = 3%.
        }
        ELSE
        {
            //NC EQ 4
            IF (spectral type SVVHNIR or SVHNIR occurs in the local histogram) THEN
            {
                Land cover class is AC/HR.
                //estimated omission error =100 * (12/35) * 9% = 34% * 9% = 3%.
                //estimated commission error (due to F/W)=100 * (3/26) * 9% = 11% * 9% = 1%.
            }
            ELSE
            {
                Land cover class is F/W.
                //estimated omission error = 100 * (3/26) * 9% = 11% * 9% =1%.
                //estimated commission error (due to AC/HR) =100 *(12/35)*9% = 34%*9% = 3%.
            }
        }       //close IF (NSC < 4) THEN …
    }           //close IF (NSC > 4) THEN …

    //Land cover class F/W:
    //        Final estimated omission error = 1% (if NSC = 4) + 3% (if NSC < 4) = 4%.
    //        Final estimated commission error = 3% (if NSC > 4) + 3% (if NSC = 4) = 6%.
    //Land cover class AC/HR:
    //        Final estimated omission error = 3% (if NSC = 4) + 3% (if NSC > 4) = 6%.
    //        Final estimated commission error = 3% (if NSC < 4) + 1% (if NSC = 4) = 4%.
}   //close IF (pixel p0 ∈ binary vegetation mask …) THEN
```

modules that are suitable for the automatic recognition of roads, buildings, and impervious surfaces in VHR images, which is in agreement with the original works by Shackelford and Davis.

ACKNOWLEDGMENT

The authors would like to thank the Editor-in-Chief, the Associate Editor, and anonymous reviewers for their patience

in reviewing this paper and for their invaluable suggestions in improving this paper.

REFERENCES

[1] G. Gutman, A. C. Janetos, C. O. Justice, E. F. Moran, and J. F. Mustard, *Land Change Science*. Dordrecht, The Netherlands: Kluwer, 2004.

[2] GEO, GEO Announces Free and Unrestricted Access to Full Landsat Archive2008.

[3] O. Sjahputera, C. H. Davis, B. Claywell, N. J. Hudson, J. M. Keller, M. G. Vincent, Y. Li, M. Klaric, and C. R. Shyu, "GeoCDX: An automated change detection and exploitation system for high resolution satellite imagery," in *Proc. IGARSS*, Boston, MA, 2008, pp. V-467–V-470.

[4] GEO, "GEO 2007-2009 Work Plan: Toward Convergence," 2008. [Online]. Available: http://earthobservations.org

[5] GEO, *The Global Earth Observation System of Systems (GEOSS) 10-Year Implementation Plan*, adopted Feb. 16, 2005. [Online]. Available: http://www.earthobservations.org/docs/10-Year%20Implementation%20Plan.pdf

[6] [Online]. Available: http://www.esa.int/esaLP/SEMOBSO4KKF_LPgmes_0.html

[7] [Online]. Available: http://www.gmes.info

[8] G. J. Hay and G. Castilla, "Object-based image analysis: Strengths, weaknesses, opportunities and threats (SWOT)," in *Proc. 1st Int. Conf. OBIA*, Salzburg, Austria, 2006.

[9] D. I. GmbH, "eCognition User Guide 4," Definiens Imaging GmbH 2004.

[10] A. Pekkarinen, L. Reithmaier, and P. Strobl, "Pan-European forest/non-forest mapping with Landsat ETM+ and CORINE Land Cover 2000 data," *ISPRS J. Photogramm. Remote Sens.*, vol. 64, no. 2, pp. 171–183, Mar. 2009.

[11] T. Esch, M. Thiel, M. Bock, A. Roth, and S. Dech, "Improvement of image segmentation accuracy based on multiscale optimization procedure," *IEEE Geosci. Remote Sens. Lett.*, vol. 5, no. 3, pp. 463–467, Jul. 2008.

[12] P. Zamperoni, "Plus ça va, moins ça va," *Pattern Recognit. Lett.*, vol. 17, no. 7, pp. 671–677, Jun. 1996.

[13] P. Corcoran and A. Winstanley, "Using texture to tackle the problem of 1456 scale in land-cover classification," in *Object-Based Image Analysis Spatial Concepts for Knowledged-Driven Remote Sensing Applications*, T. Blaschke, S. Lang, and G. J. Hay, Eds. Berlin, Heidelberg: Springer, 2008, pp. 113–132.

[14] M. Page-Jones, *The Practical Guide to Structured Systems Design*. Englewood Cliffs, NJ: Prentice-Hall, 1988.

[15] V. Cherkassky and F. Mulier, *Learning From Data: Concepts, Theory, and Methods*. New York: Wiley, 1998.

[16] A. Baraldi and F. Parmiggiani, "Combined detection of intensity and chromatic contours in color images," *Opt. Eng.*, vol. 35, no. 5, pp. 1413–1439, May 1996.

[17] T. Matsuyama and V. S.-S. Hwang, *SIGMA—A Knowledge-Based Aerial Image Understanding System*. New York: Plenum, 1990.

[18] C. Hudelot, J. Atif, and I. Bloch, "Fuzzy spatial relation ontology for image interpretation," *Fuzzy Sets Syst.*, vol. 159, no. 15, pp. 1929–1951, Aug. 2008.

[19] A. K. Shackelford and C. H. Davis, "A combined fuzzy pixel-based and object-based approach for classification of high-resolution multispectral data over urban areas," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 10, pp. 2354–2363, Oct. 2003.

[20] A. K. Shackelford and C. H. Davis, "A hierarchical fuzzy classification approach for high-resolution multispectral data over urban areas," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 9, pp. 1920–1932, Sep. 2003.

[21] A. K. Shackelford and C. H. Davis, "Fully automated road network extraction from high-resolution satellite multispectral imagery," in *Proc. IGARSS*, 2003, pp. 461–463.

[22] A. Baraldi, L. Durieux, D. Simonetti, G. Conchedda, F. Holecz, and P. Blonda, "Automatic spectral rule-based preliminary classification of radiometrically calibrated SPOT-4/-5/IRS, AVHRR/MSG, AATSR, IKONOS/QuickBird/OrbView/GeoEye and DMC/SPOT-1/-2 imagery—Part I: System design and implementation," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 3, pp. 1299–1325, Mar. 2010.

[23] D. Marr, *Vision*. New York: Freeman, 1982.

[24] M. Sonka, V. Hlavac, and R. Boyle, *Image Processing, Analysis and Machine Vision*. London, U.K.: Chapman & Hall, 1994.

[25] A. Baraldi, L. Durieux, D. Simonetti, G. Conchedda, F. Holecz, and P. Blonda, "Automatic spectral rule-based preliminary classification of radiometrically calibrated SPOT-4/-5/IRS, AVHRR/MSG, AATSR, IKONOS/QuickBird/OrbView/GeoEye and DMC/SPOT-1/-2 imagery—Part II: Classification accuracy assessment," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 3, pp. 1326–1354, Mar. 2010.

[26] A. Baraldi, V. Puzzolo, P. Blonda, L. Bruzzone, and C. Tarantino, "Automatic spectral rule-based preliminary mapping of calibrated Landsat TM and ETM+ images," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 9, pp. 2563–2586, Sep. 2006.

[27] A. Baraldi, M. Gironda, and D. Simonetti, "Operational two-stage stratified topographic correction of spaceborne multi-spectral imagery employing an automatic spectral rule-based decision-tree preliminary classifier," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 1, pp. 112–146, Jan. 2010.

[28] Q. Yu and D. A. Clausi, "SAR sea-ice image analysis based on iterative region growing using semantics," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 12, pp. 3919–3931, Dec. 2007.

[29] P. H. Swain and S. M. Davis, *Remote Sensing: The Quantitative Approach*. New York: McGraw-Hill, 1978.

[30] F. Wang, "Fuzzy supervised classification of remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 28, no. 2, pp. 194–201, Mar. 1990.

[31] K. M. Adeney and M. Korenberg, "On the use of separable Volterra networks to model discrete-time Volterra systems," *IEEE Trans. Neural Netw.*, vol. 12, no. 1, pp. 174–175, Jan. 2001.

[32] A. Bharath and M. Petrou, *Next Generation Artificial Vision Systems—Reverse Engineering the Human Visual System*. Boston, MA: Artech House, 2008.

[33] T. Lindeberg, "Detecting salient blob-like image structures and their scales with a scale-space primal sketch: A method for focus-of-attention," *Int. J. Comput. Vis.*, vol. 11, no. 3, pp. 283–318, Dec. 1993.

[34] C. Carson, S. Belongie, H. Greenspan, and J. Malik, "Region-based image querying," in *Proc. Int. Workshop Content-Based Access Image Video Libraries*, 1997, pp. 42–49.

[35] J. Yang and R. S. Wang, "Classified road detection from satellite images based on perceptual organization," *Int. J. Remote Sens.*, vol. 28, no. 20, pp. 4653–4669, Oct. 2007.

[36] D. C. Burr and M. C. Morrone, "A nonlinear model of feature detection," in *Nonlinear Vision: Determination of Neural Receptive Fields, Functions, and Networks*, R. B. Pinter and B. Nabet, Eds. Boca Raton, FL: CRC Press, 1992, pp. 309–327.

[37] L. Delves, R. Wilkinson, C. Oliver, and R. White, "Comparing the performance of SAR image segmentation algorithms," *Int. J. Remote Sens.*, vol. 13, no. 11, pp. 2121–2149, 1992.

[38] M. Petrou and P. Sevilla, *Image Processing: Dealing With Texture*. New York: Wiley, 2006.

[39] M. Nagao and T. Matsuyama, *A Structural Analysis of Complex Aerial Photographs*. New York: Plenum, 1980.

[40] C. M. Bishop, *Neural Networks for Pattern Recognition*. Oxford, U.K.: Clarendon, 1995.

[41] D. Chakraborty and N. R. Pal, "A neuro-fuzzy scheme for simultaneous feature selection and fuzzy rule-based classification," *IEEE Trans. Neural Netw.*, vol. 15, no. 1, pp. 110–123, Jan. 2004.

[42] A. Baraldi, "Impact of radiometric calibration and specifications of spaceborne optical imaging sensors on the development of operational automatic remote sensing image understanding systems," *IEEE J. Sel. Topics Appl. Earth Obs. Remote Sens.*, vol. 2, no. 2, pp. 104–134, Jun. 2009.

[43] R. G. Congalton and K. Green, *Assessing the Accuracy of Remotely Sensed Data*. Boca Raton, FL: Lewis Publishers, 1999.

[44] C. Mason and E. R. Kandel, "Central visual pathways," in *Principles of Neural Science*, E. Kandel and J. Schwartz, Eds. Norwalk, CT: Appleton and Lange, 1991, pp. 420–439.

[45] P. Mather, *Computer Processing of Remotely-Sensed Images—An Introduction*. Chichester, U.K.: Wiley, 1994.

[46] [Online]. Available: http://www.coventry.ac.uk/ec/~nhunt/meths/strati.html

[47] K. Pakzad, J. Bückner, and S. Growe, "Knowledge based moorland interpretation using a hybrid system for image analysis," in *Proc. ISPRS Conf.*, Munich, Germany, Sep. 8–10, 1999.

[48] R. Richter, Atmospheric/Topographic Correction for Satellite Imagery—ATCOR-2/3 User Guide, Wessling, Germany: DLR2006. [Online]. Available: http://www.geog.umontreal.ca/donnees/geo6333/atcor23_manual.pdf

[49] GEO/CEOSS, "A Quality Assurance Framework for Earth Observation," version 2.0, Sep. 2008. [Online]. Available: http://calvalportal.ceos.org/CalValPortal/showQA4EO.do?section=qa4eoIntro

[50] A. Strahler, D. Muchoney, J. Borak, M. Friedl, S. Gopal, E. Lambin, and A. Moody, "Modis Land Cover Product Algorithm Theoretical Basis Document (ATBD)—version 5.0," MODIS Science Team. [Online]. Available: http://modis.gsfc.nasa.gov/data/atbd/atbd_mod12.pdf

[51] Image2000 and CORINE Land Cover (CLC) 2000. *Eur. Comm. Joint Res. Center*, EUR 21757, Ispra, Italy, 2005.
[52] T. Lillesand and R. Kiefer, *Remote Sensing and Image Interpretation*. New York: Wiley, 1994.
[53] M. Herold, C. Woodcock, A. Di Gregorio, P. Mayaux, A. S. Belward, J. Latham, and C. Schmullius, "A joint initiative for harmonization and validation of land cover datasets," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 7, pp. 1719–1727, Jul. 2006.
[54] A. Di Gregorio and L. Jansen, "Land Cover Classification System (LCCS): Classification concepts and user manual," FAO Corporate Document Repository. [Online]: Available: http://www.fao.org/DOCREP/003/X0596E/X0596e00.htm
[55] L. Prechelt, "A quantitative study of experimental evaluations of neural network learning algorithms: Current research practice," *Neural Netw.*, vol. 9, no. 3, pp. 457–462, Apr. 1996.
[56] [Online]. Available: http://eusoils.jrc.ec.europa.eu
[57] [Online]. Available: http://www.fao.org/ag/agl/agll/key2soil.stm
[58] ENVI 4.3 User Manual, 2006.
[59] P. S. Chavez, "An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data," *Remote Sens. Environ.*, vol. 24, no. 3, pp. 459–479, Apr. 1988.
[60] G. M. Foody, "Status of land cover classification accuracy assessment," *Remote Sens. Environ.*, vol. 80, no. 1, pp. 185–201, Apr. 2002.
[61] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 12, no. 7, pp. 629–639, Jul. 1990.
[62] E. Lunetta and C. Elvidge, *Remote Sensing Change Detection: Environmental Monitoring Methods and Applications*. London, U.K.: Taylor & Francis, 1999.
[63] A. K. Jain, R. P. W. Duin, and J. Mao, "Statistical pattern recognition: A review," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 1, pp. 4–37, Jan. 2000.
[64] S. Liang, *Quantitative Remote Sensing of Land Surfaces*. Hoboken, NJ: Wiley-Interscience, 2004.
[65] J. R. G. Townshend, *The Spatial Resolving Power of Earth Resources Satellites: A Review*. Greenbelt, MD: NASA, 1980.
[66] M. J. Barnsley and S. A. W. Kay, "The relationship between sensor geometry, vegetation-canopy geometry and image variance," *Int. J. Remote Sens.*, vol. 11, no. 6, pp. 1075–1083, Jun. 1990.
[67] D. S. Kimes, "Dynamics of directional reflectance factor distributions for vegetation canopies," *Appl. Opt.*, vol. 22, no. 9, pp. 1364–1372, May 1983.
[68] T. N. Carlson and D. A. Ripley, "On the relation between NDVI, fractional vegetation cover, and leaf area index," *Remote Sens. Environ.*, vol. 62, no. 3, pp. 241–252, Dec. 1997.
[69] R. Trias-Sanz, "Texture orientation and period estimator for discriminating between forests, orchards, vineyards, and tilled fields," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 10, pp. 2755–2760, Oct. 2006.
[70] H. Anys and D.-C. He, "Evaluation of textural and multipolarization radar features for crop classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 33, no. 5, pp. 1170–1181, Sep. 1995.
[71] B. Julesz, "Texton gradients: The texton theory revisited," *Biol. Cybern.*, vol. 54, no. 4/5, pp. 245–251, Aug. 1986.
[72] J. I. Yellott, "Implications of triple correlation uniqueness for texture statistics and the Julesz conjecture," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 10, no. 5, pp. 777–793, May 1993.
[73] J. Canny, "A computational approach to edge detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-8, no. 6, pp. 679–698, Nov. 1986.
[74] A. E. Magurran, *Ecological Diversity and Its Measurement*. Princeton, NJ: Princeton Univ. Press, 1988.
[75] A. Baraldi, L. Bruzzone, and P. Blonda, "Quality assessment of classification and cluster maps without ground truth knowledge," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 4, pp. 857–873, Apr. 2005.
[76] R. Nishii and S. Eguchi, "Supervised image classification by contextual AdaBoost based on posteriors in neighborhoods," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 11, pp. 2547–2554, Nov. 2005.
[77] A. Vehtari and J. Lampinen, "Bayesian MLP neural networks for image analysis," *Pattern Recognit. Lett.*, vol. 21, no. 13/14, pp. 1183–1191, Dec. 2000.
[78] D. Lu and Q. Weng, "A survey of image classification methods and techniques for improving classification performance," *Int. J. Remote Sens.*, vol. 28, no. 5, pp. 823–870, Mar. 2007.
[79] N. R. Pal and S. K. Pal, "A review on image segmentation techniques," *Pattern Recognit.*, vol. 26, no. 9, pp. 1277–1294, Sep. 1993.

Andrea Baraldi received the Laurea (M.S.) degree in electronic engineering from the University of Bologna, Bologna, Italy, in 1989, where his master's thesis focused on the development of segmentation and classification algorithms for remotely sensed optical imagery.

From 1989 to 1990, he worked as a Research Associate with the Centro di Studio per l'Interazione Operatore–Calcolatore, National Research Council (CNR), Bologna, and served in the army at the Istituto Geografico Militare, Florence, Italy, working on satellite image classifiers and geographic information systems (GIS). From 1991 to 1993, as a Consultant with the European Space Agency–European Space Research Institute, Frascati, Italy, he worked on object-oriented applications for GIS. From December 1997 to June 1999, he was assigned a post-doctoral fellowship in artificial intelligence with the International Computer Science Institute, Berkeley, CA. From 2000 to 2002, as a Postdoctoral Researcher, he joined the Global Vegetation Monitoring Unit, Institute for Environment and Sustainability (IES), Joint Research Center (JRC), European Commission, Ispra, Italy, in the development and validation of algorithms for forest classification of radar mosaics at continental scale. From 2005 to 2009, at the IES-Spatial Data Infrastructure, JRC, he was involved with satellite optical image calibration, classification, and mosaicking at continental scale. Since his master's thesis, he has been continuing his collaboration with Istituto di Scienze dell'Atmosfera e del Clima, CNR, Bologna, and Istituto di Studi di Sistemi Intelligenti per l'Automazione, CNR, Bari, Italy. He is currently with Baraldi Consultancy in Remote Sensing, Bologna. His main interests center on image understanding, with special emphasis on the development of operational automatic hierarchical multisource multiresolution spaceborne image understanding systems that are consistent with a biological vision.

Mr. Baraldi served as an Associate Editor of the IEEE TRANSACTIONS ON NEURAL NETWORKS from 2001 to 2006.

Tom Wassenaar received the M.Sc. degree in physical geography from the University of Amsterdam, Amsterdam, The Netherlands, in 1995 and the Ph.D. degree in soil science from the Ecole Nationale Supérieure d'Agronomie, Montpellier, France, in 2001 as an EU-funded Marie Curie Research Fellow with the French agronomic research institute Institut National de la Recherche Agronomique, where his work was on remote sensing of soil surface features in a perennial crop (vineyard) environment.

He has been performing research on agroenvironmental issues, particularly with the Food and Agricultural Organization, UN, and on agriculture-related very high resolution remote sensing applications, particularly with the Agriculture Unit, Institute for the Protection and Security of the Citizen, Joint Research Center, European Commission, Ispra, Italy. He is currently with the Centre de Coopération Internationale en Recherche Agronomique pour le Développement (CIRAD), which is a French research center working with developing countries to tackle international agricultural and development issues.

Simon Kay received the B.A. Hons Geography degree, the M.Sc. degree in remote sensing, and the Ph.D. degree in geography, specializing in airborne remote sensing of agricultural crops, from University College, London, U.K..

After initially working internationally in a broad range of geomatics applications as a Consultant, he joined the European Commission's Joint Research Center (JRC) in 1995, working on agricultural subsidy monitoring, olive tree recognition and mapping, and establishment of geographic information systems technology to map farmer fields across Europe. He is currently the Acting Head of the Monitoring Agricultural Resources Unit, Institute for the Protection and Security of the Citizen, JRC, Ispra, Italy.