Quality Assessment of Preclassification Maps Generated From Spaceborne/Airborne Multispectral Images by the *Satellite Image Automatic Mapper* and *Atmospheric/Topographic Correction-Spectral Classification* Software Products: Part 1—Theory

Andrea Baraldi and Michael L. Humber

***Abstract*—In compliance with the Quality Assurance Framework for Earth Observation (QA4EO) guidelines, the goal of this paper is to provide a theoretical comparison and an experimental quality assessment of two operational (ready-for-use) expert systems (prior knowledge-based nonadaptive decision trees) for automatic near real-time preattentional classification and segmentation of spaceborne/airborne multispectral (MS) images: the Satellite Image Automatic Mapper™ (SIAM™) software product and the Spectral Classification of surface reflectance signatures (SPECL) secondary product of the Atmospheric/Topographic Correction™ (ATCOR™) commercial software toolbox. For the sake of simplicity, this paper is split into two: Part 1—Theory, presented herein, and Part 2—Experimental results, already published elsewhere. The main theoretical contribution of the present Part 1 is threefold. First, it provides the published Part 2 with an interdisciplinary terminology and a theoretical background encompassing multiple disciplines, such as philosophical hermeneutics, machine learning, artificial intelligence, computer vision, human vision, and remote sensing (RS). Second, it highlights the several degrees of novelty of the ATCOR-SPECL and SIAM deductive preliminary classifiers (preclassifiers) at the four levels of abstraction of an information processing system, namely, system design, knowledge/information representation, algorithms, and implementation. Third, the present Part 1 requires the experimental Part 2 to collect a minimum set of complementary statistically independent metrological quality indicators (QIs) of operativeness (QIOs), in compliance with the QA4EO guidelines and the principles of statistics. In particular, sample QIs are required to be: 1) statistically significant, i.e., provided with a degree of uncertainty in measurement; and 2) statistically valid (consistent), i.e., representative of the entire population being sampled, which requires the implementation of a probability sampling protocol. Largely overlooked by the RS community, these sample QI requirements are almost never satisfied in the RS common practice. As a consequence, to date, QIOs of existing RS image understanding systems (RS-IUSs), including thematic map accuracy, remain largely unknown in statistical terms. The conclusion of the present Part 1 is that the proposed comparison of the two alternative ATCOR-SPECL and SIAM prior knowledge-based preclassifiers in operating mode, accomplished in the Part 2, can be considered appropriate, well-timed, and of potential interest to a large portion of the RS readership.**

***Index Terms*—Attentive vision, degree of uncertainty in measurement, land cover classification taxonomy, preattentive vision, preliminary classification, probability sampling, quality indicator (QI), radiometric calibration, spectral category, spectral mixture analysis.**

Manuscript received January 05, 2014; accepted August 06, 2014. This work was supported by the National Aeronautics and Space Administration under Grant NNX07AV19G issued through the Earth Science Division of the Science Mission Directorate.

A. Baraldi was with the University of Maryland, College Park, MD 20742 USA. He is now with Baraldi Consultancy in Remote Sensing, Modena 41121, Italy (e-mail: andrea6311@gmail.com).

M. L. Humber is with the Department of Geographical Sciences, University of Maryland, College Park, MD 20742 USA (e-mail: mhumber@umd.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JSTARS.2014.2349932

I. Introduction

ONE VISIONARY goal of the remote sensing (RS) community is to develop information processing systems capable of automatically transforming, without user interactions, large-scale multisource multiresolution Earth observation (EO) image databases into "operational, comprehensive, and timely knowledge/information products" [1]–[3], at spatial extents ranging from local to global [4]. The Quality Assurance Framework for EO (QA4EO) guidelines [2], [3], conceived by the international Group on EOs (GEO)-Committee on EO Satellites (CEOS), comprise an extensive formulation of this ambitious goal. For example, the ongoing GEO Global EO System of Systems (GEOSS) implementation plan for years 2005–2015 incorporates the QA4EO guidelines to build a global public infrastructure that allows "the provision of and access to the Right (geospatial) Information, in the Right Format, at the Right Time, to the Right People, to Make the Right Decisions" [1].

To pave the way for the design and implementation of a novel generation of automatic RS image understanding systems (RS-IUSs) in compliance with the QA4EO guidelines [2], [3], this paper provides a theoretical comparison and an experimental quality assessment of two operational (ready-for-use) expert systems (prior knowledge-based nonadaptive decision trees) for automatic near real-time preliminary classification (preclassification [5]) and segmentation of spaceborne/airborne EO multispectral (MS) images: the spectral classification of surface reflectance signatures (SPECL) software module and the Satellite Image Automatic Mapper (SIAM) software product. The former is implemented as a nonvalidated secondary product within the popular Atmospheric/Topographic Correction (ATCOR)-2/3/4 commercial software toolbox [6]–[9]. The latter has been presented in recent years in the RS literature [10]–[19], where enough information is provided for the SIAM implementation to be reproduced [11], [17].

Rather than being considered as standalone software products, the two alternative ATCOR-SPECL and SIAM expert systems for automatic near real-time preclassification and segmentation of multisource MS images are eligible for use in the preattentive vision first stage of a novel generation of automatic *hybrid* (combined deductive and inductive) RS-IUS implementations [10]–[20].

For the sake of simplicity, this paper is split into two: the Part 1—Theory, presented herein, and the Part 2—Experimental results, already published elsewhere [20]. The main theoretical contribution of the present Part 1 is threefold. First, it provides the Part 2 with an interdisciplinary terminology and a theoretical background encompassing multiple disciplines, such as philosophical hermeneutics, machine learning, artificial intelligence, computer vision, human vision, and RS. Hence, Part 1 is provided with a relevant survey value. Second, it highlights the relevant degrees of novelty of the ATCOR-SPECL and SIAM prior knowledge-based preclassifiers at the four levels of abstraction of an information processing system, namely, system design, knowledge/information representation, algorithms, and implementation. Third, the present Part 1 requires the experimental Part 2 to collect a minimum set of complementary independent metrological/statistically-based quality indicators (QIs) of operativeness (QIOs), in compliance with the QA4EO guidelines and the principles of statistics. In particular, sample QIs are required to be: 1) statistically significant, i.e., provided with a degree of uncertainty in measurement and 2) statistically valid (consistent), i.e., representative of the entire population being sampled, which requires the implementation of a probability sampling protocol. Largely overlooked by the RS community, these sample QI requirements are almost never satisfied in the RS common practice. As a consequence, to date, QIOs of existing RS-IUSs, including thematic map accuracy, remain largely unknown in statistical terms. The conclusion of the present Part 1 is that the proposed comparison of the two alternative ATCOR-SPECL and SIAM prior knowledge-based preclassifiers in operating mode, accomplished in the Part 2, can be considered appropriate, well-timed, and of potential interest to a large portion of the RS readership.

The rest of the present Part 1 is organized as follows. Section II presents an interdisciplinary terminology and a theoretical background useful for the understanding of the experimental Part 2. Problem recognition and opportunity identification are discussed in Section III. In Section IV, the two alternative ATCOR-SPECL and SIAM preclassification expert systems are compared at the four levels of abstraction of an information processing system. Conclusion of this theoretical contribution is reported in Section V.

II. Interdisciplinary Terminology and Problem Background

According to Section I, the goal of the experimental Part 2 of this paper, published elsewhere [20], is to pursue a statistically significant and statistically consistent quality assessment of the ATCOR-SPECL and SIAM deductive preclassification software products in operating mode, eligible for use in the preattentive vision first stage of a hybrid RS-IUS architecture [20]. Introduced by Section I, terms such as “statistically significant” QI, “statistically consistent” probability sampling, “QIOs of an information processing system in operating mode,” “quality assessment of a preclassification map,” “deductive preclassification,” “preattentive/attentive vision,” “deductive/inductive/hybrid inference,” and “data/information/knowledge” are defined explicitly and unambiguously in this section, based on a multidisciplinary approach. To be employed in the rest of the present Part 1 and in the Part 2, the proposed interdisciplinary terminology provides this paper with a significant survey value.

A. Quantitative and Qualitative Concepts of Information

Philosophical hermeneutics refers to the theory of knowledge and the practice, art or science of (text) interpretation and explanation. According to philosophical hermeneutics [21], [22], the impact upon computer science, information technology (IT), artificial intelligence and machine learning of existing different quantitative and qualitative concepts of information, embedded in more or less explicit information theories, appears largely underestimated. This means that fundamental questions—like: When do (subsymbolic) data become (symbolic) information [23]? When does vision go symbolic [5]? Should traditional information retrieval be called document retrieval [21], [22]?—appear largely overlooked and, as a consequence, far from being answered.

In accordance with philosophical hermeneutics, the fundamental concepts of *numerical data*, *quantitative information*, *qualitative information* and *knowledge* are defined hereafter [21], [22].

1) Numerical data, sensory data, quantitative data, observational data are considered synonyms of “true facts” [24]. *Sensory data are provided, per se, with no semantics at all* [23], i.e., observational data are always subsymbolic (unlabeled).
2) Subsymbolic, quantitative, unequivocal “*information-as-thing*” is, according to the Shannon theory of communication [25], an object or a thing (e.g., number of bits and number of words in a document) irrespective of its meaning. This makes the information exchange between a sender and a receiver unequivocal (context independent) and, therefore, easier to deal with than when meaning is involved in the communication process [18], [19], [21], [22].
3) Symbolic, qualitative, equivocal “*information-as-(an intepretation)process*,” i.e., information as interpreted data, is, in the words of philosophical hermeneutics, symbolic information always related to “a receiver’s beliefs,

desires and background knowledge" [21], [22]: the meaning of a message is always context-dependent, depending on (changing with) the inquirer (user, knower, receiver, cognitive agent) in charge of the message interpretation. For example, Adams *et al.* underline that land cover (LC) "class names are selected to have significance to an observer in the field and in the context of a given study" [26].

4) "Knowledge" is strictly related to the concept of "*information-as-(an intepretation)process*," such that "there is no knowledge without both an object of knowledge and a knowing subject." [21], [22]. Hence, "information-as-(an interpretation)process" and "knowledge" can be considered as synonyms. A well-known example of equivocal (subjective, context-dependent) interpretation process is the so-called "fusion of ontologies" or "fusion of thematic map legends" [21], [22], occurring when two thematic maps of the same geographic area, but featuring different map legends, must be compared. In other words, it is reasonable to expect that two independent domain experts required to harmonize (reconcile) two thematic map legends may fulfill their (inherently equivocal) interpretation processes with different inter-vocabulary mapping functions.

Noteworthy, *the complementary concepts of information-as-(an interpretation)process and information-as-thing apply one-to-one to the dual concepts of (equivocal, qualitative, symbolic) categorical (nominal) variables and (unequivocal, quantitative, subsymbolic) continuous/discrete scalar/vector variables* (e.g., biophysical variables, such as leaf area index and biomass), *to be estimated from sensory data* [18], [19], [47]. To conclude, the following terms can be considered as nontrivial synonyms.

1) Symbolic, semantic, cognitive, categorical, ordinal, nominal, qualitative, subjective, equivocal. For example, (discrete and symbolic) categorical variable.
2) Subsymbolic, sensory, numerical, nonsemantic, quantitative, objective, unequivocal. For example, (subsymbolic) continuous or discrete sensory variable.

For example, according to the terminology proposed herein, the two ATCOR-SPECL and SIAM prior knowledge-based preclassifiers, to be assessed and compared in the Part 2 [20], automatically transform (subsymbolic quantitative) MS images (2-D data) into a (symbolic qualitative) categorical variable, whose values belong to a discrete and finite legend of (semantic) concepts.

B. Inductive, Deductive, and Hybrid Inference Systems, Either Subsymbolic or Symbolic, Investigated by the Machine Learning, Artificial Intelligence, and RS Disciplines

This section introduces expressions like inductive, deductive and hybrid inference system, either subsymbolic or symbolic (refer to Section II-A), depending on whether the inference system deals with, respectively, subsymbolic variables, either continuous or discrete, or (symbolic and discrete) categorical (nominal) variables. The specialization capability of this terminology is far superior to that of expressions traditionally used or misused by the RS community, such as supervised or unsupervised data learning. For example, an expression such as "unsupervised classification" is widely adopted by the RS community to mean either "unsupervised data clustering" or "automatic classification," e.g., see [27] and [28]. Unfortunately, according to the machine learning literature, this expression is a typical contradiction of terms because: 1) "unsupervised," e.g., unsupervised data, refers to "unlabeled," e.g., unlabeled data, rather than "without user's supervision," i.e., "unsupervised" does not mean "automatic" and 2) sensory data are provided with no semantics at all (refer to Section II-A), i.e., observational data are always, *per se*, unsupervised (unlabeled), while, by definition, classified data are always supervised (labeled) data, where data labels belong to a discrete and finite taxonomy of (semantic) concepts [23], [24], [29].

Hereafter, the concepts of inductive, deductive and hybrid inference system, either subsymbolic or symbolic, are discussed in detail.

There are two classical types of inference (learning), known as: 1) *induction*, progressing from particular cases (e.g., true facts and training data samples) to a general estimated dependency or model, and 2) *deduction*, progressing from a general model (e.g., a physical model-based equation) to particular cases (e.g., output values) [24]. Inductive inference is the basis of the machine learning discipline [24], [29]. Deductive inference is the main focus of interest of traditional artificial intelligence [24], [29]–[31].

The following terms are nontrivial synonyms of deductive inference and become interchangeable in the rest of this work [18], [19]: (subsymbolic or symbolic) deductive inference, deductive learning, top-down inference system, coarse-to-fine inference, driven-by-knowledge inference, learning-by-rules, physical model, prior knowledge-based decision system, rule-based system, expert system, syntactic inference, and syntactic pattern recognition.

The following terms are nontrivial synonyms of inductive inference [18], [19]: (subsymbolic or symbolic) inductive inference, inductive learning from either labeled (supervised) or unlabeled (unsupervised) data, bottom-up inference, fine-to-coarse inference, driven-without-knowledge (knowledge-free) inference, learning-from-examples, statistical model.

For the sake of completeness, some well-known examples of inductive and deductive inference systems, presented in the computer vision, machine learning and/or RS literature, are listed as follows.

1) In the computer vision literature, image segmentation algorithms are typical examples of subsymbolic inductive inference systems for unlabeled data learning [32]–[36].
2) In the machine learning literature, unsupervised (unlabeled) data learning algorithms are either vector data quantizers (e.g., the well-known k-means data quantization algorithm, improperly called k-means data clustering algorithm), probability density function estimators or unlabeled data clustering algorithms [15], [24], [29], [37]–[40]. Inductive supervised (labeled) data learning systems are either: 1) symbolic (classifiers), e.g., artificial neural network classifiers, support vector machine

classifiers [41], nearest-neighbor classifiers, adaptive decision-tree classifiers, and radial basis function networks for classification [24], [29] or 2) subsymbolic, suitable for function regression, e.g., radial basis function networks for function regression [24], [29].

3) In the RS literature [24], [29], a typical example of subsymbolic inductive inference system is principal component analysis; a popular example of subsymbolic deductive inference system is tasseled cap transformation.

The machine learning literature clearly acknowledges that all inductive data learning problems are inherently ill-posed in the Hadamard sense [42]. According to Hadamard, mathematical or statistical models of physical phenomena are defined as well-posed (respectively, ill-posed) when they satisfy (respectively, do not satisfy at least one of) the following requirements [42]: 1) a solution exists, 2) the solution is unique, and 3) the solution's behavior hardly changes when there is a slight change in the initial condition. In the words of Mulier and Cherkassky: "*induction amounts to forming generalizations from particular true facts. This is an inherently difficult (ill-posed) problem and its solution requires a priori knowledge in addition to data*" [24] (p. 39). Hence, to become better posed (conditioned) for numerical treatment, any inductive data learning algorithm requires an *a priori* knowledge base (deductive inference approach) to avoid starting from scratch when looking at input sensory data [10]–[19]. This conclusion complies with the well-known statistical principle of *stratification*, equivalent to the divide-and-conquer (*dividi et impera*) problem solving approach [29], to be enforced upon statistical systems. The advantage of a stratified statistical system is that it "*will always achieve greater precision (than its nonstratified counterpart), 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*" [43].

On one hand, well-known limitations of statistical (bottom-up inference) systems in common practice are that they are inherently semiautomatic and site-specific [18], [45]. On the other hand, typical drawbacks of physical (top-down inference) models are that [18]: 1) in general, it takes a long time for human experts to learn physical laws of the real-world-through-time and tune physical models, 2) physical models suffer from an intrinsic lack of flexibility, i.e., decision rules do not adapt to changes in the input data format and users' needs, hence their knowledge base may soon become obsolete, and 3) physical models suffer from an intrinsic lack of scalability, in particular rule-based systems are impractical for complex problems [30].

There is an ongoing multidisciplinary debate about a claimed inadequacy of scientific disciplines such as computer vision, artificial intelligence, and machine learning, whose origins date back to the late 1950s, in the provision of operational solutions to their ambitious cognitive objectives [23], [44]. This claim may mean that, if they are not combined, inductive and deductive inference approaches show intrinsic weaknesses in operational use, irrespective of implementation [18]. As a consequence, to outperform existing deductive and inductive inference systems whose drawbacks are well known, a novel trend in recent literature aims at developing hybrid inference systems for retrieval of subsymbolic variables (e.g., leaf area index, LAI) or symbolic variables (e.g., LC and LC change (LCC) classes) from sensory data (e.g., optical imagery) [45]–[48]. By definition, *hybrid inference systems, either subsymbolic or symbolic, combine both statistical and physical models to take advantage of the unique features of each and overcome their shortcomings* [46], [47]. For example, in the foreword of the seminal book by Nagao and Matsuyama [47], published in 1980 (oldies, but goldies), it is written: "The work described here is a deep *unification and synthesis of the two fundamental approaches to pattern recognition: numerical (also known as 'statistical') and structural ('linguistic,' 'syntactic')*."

Noteworthy, physical model-based inference systems as well as hybrid models require as input observational data provided with a physical meaning, i.e., sensory data provided with a physical unit of measure, e.g., RS imagery radiometrically calibrated into top-of-atmosphere (TOA) radiance or TOA reflectance values [10]. On the other hand, statistical systems can be input with any sort of numerical data, irrespective of their physical meaning, if any. This is tantamount to saying that, whereas *dimensionless sensory data, provided with no physical unit of measure, are eligible for use as input to statistical models exclusively*, on the contrary, *numerical data provided with a physical unit of measure can be input to both physical and statistical models.*

For the sake of completeness, let us review some additional examples of inductive, deductive and hybrid RS-IUS instances proposed in recent years in the RS literature. A large family of one-stage one-pass (noniterative) prior knowledge-based (static, nonadaptive to input data) decision-tree (pre)classifiers (symbolic expert systems) has been proposed, starting from the 1970 s, as a legacy of traditional artificial intelligence [49], [50], [51]–[54]. For example, in [50] (p. 4176), a one-stage physical model-based RS-IUS, see Fig. 1(a), consists of a hierarchy of five pixel-specific prior knowledge-based spectral rules proposed to detect six land surface types, namely, "vegetated lands," "nonvegetated lands," "snow/ice," "water bodies," "clouds," and "cloud shadows," in radiometrically calibrated 500 m resolution moderate resolution imaging spectroradiometer (MODIS) images. In 30 m resolution Landsat images, a one-stage deductive RS-IUS, consisting of a hierarchy of per-pixel prior knowledge-based spectral rules, detects LC classes "water," "coniferous forest," "deciduous forest," "agricultural areas," "grassland," "urban areas," and "roads" [52]. In recent years, prior knowledge-based decision-tree classifiers are employed per image-object at an attentive vision second stage, in series with an inductive image segmentation first stage, like in the popular two-stage noniterative Geographic Object-Based Image Analysis (GEOBIA) system architecture, see Fig. 1(b), and in the three-stage iterative Geographic Object-Observation Image Analysis (GEOOIA) system design [32]–[34], [55]–[60]. The former is a special case of the latter, i.e., GEOBIA $\subseteq$ GEOOIA, where both GEOBIA and GEOOIA share a statistical model-based subsymbolic image segmentation first stage. Alternative to GEOBIA/GEOOIA systems, an original two-stage hybrid RS-IUS architecture is proposed by Shackelford and Davis [61], [62]. It comprises an image-object-based expert system for second-stage decision-tree classification in series with a first-stage pixel-based

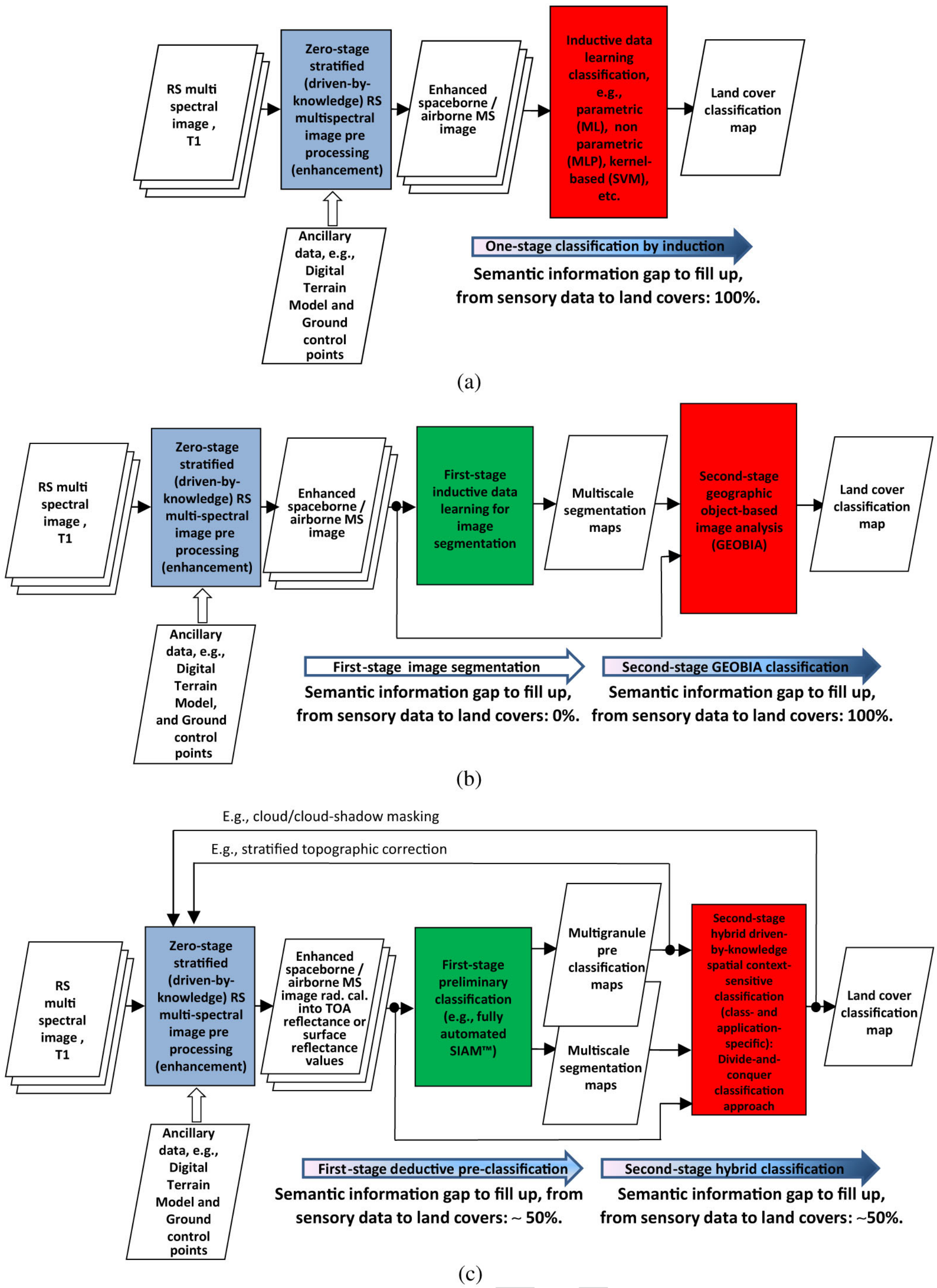

Fig. 1. (a) Top: Traditional one-stage RS-IUS architecture. 100% of the semantic information gap from sensory data to LC classes is filled up in one step. (b) Middle. Traditional two-stage noniterative GEOBIA design. 100% of the semantic information gap from sensory data to LC classes is filled up in the segment-based image classification second stage, in series with the subsymbolic inductive-data-learning image segmentation first stage. (c) Bottom. Novel three-stage hybrid RS-IUS design. Approximately, 50% of the semantic information gap from sensory data to LC classes is filled up in the automatic deductive preclassification first stage [80].

statistical preclassifier, implemented as a traditional plug-in (nonadaptive to input data) pixel-based maximum likelihood (ML) classifier. In this scenario, the ATCOR-SPECL [6]–[9] and SIAM [10]–[19] software products, to be assessed and compared in the Part 2 of this paper [20], are, to the best of these authors' knowledge, the first examples of prior knowledge-based decision-tree preclassifiers in operating mode eligible for use at the preattentive vision first stage of a hybrid RS-IUS architecture, see Fig. 1(c). Noteworthy, the hybrid RS-IUS architecture shown in Fig. 1(c) is alternative to both the two-stage hybrid RS-IUS architecture proposed by Shackelford and Davis [61], [62] and the GEOBIA/GEOOIA system architecture shown in Fig. 1(b). To summarize, whereas prior knowledge-based decision-tree classifiers have been traditionally employed in one-stage RS-IUSs [see Fig. 1(a)] or at the attentive vision second stage of two-stage hybrid

RS-IUSs, whose first stage consists of either a subsymbolic statistical system, like in GEOBIA/GEOOIA systems, see Fig. 1(b), or a semisymbolic plug-in statistical system, like in the Shackelford and Davis RS-IUS architecture [61], [62], the degree of novelty of the ATCOR-SPECL and SIAM prior knowledge-based preclassifiers is to provide a multistage hybrid RS-IUS architecture with an injection of prior knowledge right at the level of the preattentive vision first stage [10]–[19], see Fig. 1(c) [20]. Additional examples of hybrid inference systems for RS image classification are those proposed by Matsuyama *et al.* in [46], [47], as well as the popular Landsat-7 Enhanced Thematic Mapper $(\mathrm{ETM}) + \mathrm{Automated}$ Cloud-Cover Assessment (ACCA) algorithm. In the ACCA algorithm, first, a per-pixel (context-independent) physical model-based decision rule set is applied to a radiometrically calibrated Landsat image to detect pixels considered as cloud candidates. Second, to remove small holes in cloud segments, a bottom-up (data-driven) context-sensitive aggregation and filling algorithm is applied in the (2-D) image domain to pixels considered as noncloud candidates at step one [63] (p. 1183).

C. Human and Computer Vision

In the words of Iqbal and Aggarwal: "frequently, no claim is made about the pertinence or adequacy of the digital models as embodied by computer algorithms to the proper model of human visual perception... This enigmatic situation arises because research and development in computer vision is often considered quite separate from research into the functioning of human vision. A fact that is generally ignored is that *biological vision is currently the only measure of the incompleteness of the current stage of computer vision, and illustrates that the problem is still open to solution*" [64].

According to this quote, human vision should be considered the gold standard (reference baseline) of the computer vision discipline, which incorporates RS image understanding as a special case. Unfortunately, the great majority of the RS community does not appear to consider biological vision as a reference baseline. In addition, relationships between the RS and computer vision communities appear weak too, the latter community considering the expertise of the former not very advanced, because traditional RS image understanding is pixel-based, where spatial (contextual) information is ignored. As a result of this lack of interdisciplinary communication, the RS community tends to underestimate the complexity of vision in general and RS image understanding in particular.

In the rest of this paper, including the experimental Part 2 [20], RS image understanding (classification, mapping) is conceived as a subset of computer vision, where human vision is adopted as a reference standard, to compare the ATCOR-SPECL and SIAM software products as alternative implementations of a prior knowledge-based preclassification first stage in a hybrid RS-IUS architecture [10]–[20] (refer to Section I). Hence, this paper complies with the aforementioned thesis by Iqbal and Aggarwal [64], but is in contrast with the majority of the RS literature, where links to computer vision and human vision disciplines are absent.

In this section, basic principles of human vision, which comprises a preattentive vision first stage and an attentive vision second stage, are briefly described [5], [46].

1) Goal of a (Biological or Artificial) Vision System: A (human or computer) visual system is a (biological or artificial) IUS suitable to provide plausible (multiple) symbolic description(s) of a 3-D scene, located in the (4-D) world-through-time, as it is observed by a (2-D) imaging sensor at a given acquisition time. The information gap between a subsymbolic (2-D) image and a symbolic (3-D) scene can be filled by conjectures that map subsymbolic image features (e.g., image-objects or, vice versa, image-contours) into symbolic classes of 4-D objects-through-time (4-D concepts-through-time, e.g., buildings and roads) belonging to the so-called preexisting (4-D) *world model* [46], [65]. A world model, also called *world ontology*, can be graphically represented as a semantic network consisting of: 1) classes of 4-D objects-through-time as nodes and 2) inter-concept relations as arcs between nodes, namely: (I) spatial relations, either topological (e.g., adjacency and inclusion) or nontopological (e.g., distance and in-between angle), (II) temporal relations and/or (III) nonspatiotemporal relations (e.g., part-of and subset-of) [18], [19], [46], [55], [66].

In terms of computational theory, the problem of image understanding (vision), from subsymbolic (2-D) imagery to symbolic description(s) of the (3-D) scene of the (4-D) world observed at a given time, belongs to the class of symbolic inductive data learning problems [24] (from sensory data to models, refer to Section II-B). As such, it is inherently ill-posed in the Hadamard sense [42] and, consequently, very difficult to solve, due to the combination of the two following *qualitative* and *quantitative information gaps* to be filled (refer to Section II-A) [18], [19], [46]: 1) The well-known *(semantic) information gap* between continuous subsymbolic sensory sensations and discrete symbolic (semantic, linguistic) persistent (stable) percepts (concepts), which has been thoroughly investigated in both philosophy and psychophysical studies of perception. In practice, "we are always seeing objects we have never seen before at the sensation level, while we perceive familiar objects everywhere at the perception level" [46]. 2) The intrinsic insufficiency of image features, namely, 0-D points, 1-D lines (e.g., contours) and 2-D polygons (image-objects), in the reconstruction of an observed (3-D) scene, due to *data dimensionality reduction* which causes, e.g., occlusion phenomena.

2) Processing Elements and Modular Structure of the Human Visual System: In mammals, a vision system accomplishes a preattentive vision first phase and an attentive vision second phase, summarized as follows.

1) Preattentive (low-level) vision extracts picture primitives based on general-purpose image processing criteria independent of the scene under analysis. It acts in parallel on the entire image as a rapid (< 50 ms) scanning system to detect variations in simple visual properties [67]–[69]. In the primary visual cortex (PVC, or area 17 of the visual cortex, or V1), single opponent and double opponent color cells are called Type I and Type II, respectively, by Wiesel and Hubel [72] (examples of Type I and Type II receptive fields can be found in [73]). Receptive fields that are spatially opponent, but not color

opponent, are called Type III [73]. Layers of PVC are vertically organized into blobs and interblob areas. The same single-opponent cells are thought to provide, in parallel, color contrast information to cells in the blobs, and achromatic contrast information to cells in the interblob regions. The visual cells heavily concentrated in cortical blobs are double-opponent cells. In the interblob areas, cortical cells belong to the hierarchy composed of simple- and complex-cell categories. A major difference between simple- and complex-cells is that the former are quasilinear while the latter exhibit a clear second-order squaring nonlinearity [98]. A regular sequence of hypercolumns is repeated over the surface of PVC, each hypercolumn occupying an area of about 1 mm^2. This repeating organization constitutes the modular structure of PVC, such that every axis of orientation, whose gradations of orientation are around 10° [67] to 15° [70], [71], is represented for every retinal position at at least four spatial scales of analysis [99]. In each hypercolumn, there are end-stopped cells, in addition to simple- and complex-cells [100]. While simple- and complex-cells are thought to accomplish line and edge extraction, end-stopped cells respond to image singularities, such as line/edge crossings, vertices of image-objects, and end-points of line segments [101].

2) Attentive (high-level) vision operates as a careful scanning system employing a focus of attention mechanism based on end-stopped cells [100], [101]. Scene subsets, corresponding to a narrow aperture of attention, are observed in sequence and each step is examined quickly (20–80 ms) [67]–[69].

It is worth noting that human achromatic vision is nearly as effective as human chromatic vision in detecting forms and accomplishing image interpretation. On an *a posteriori* basis, this observation has two important implications. First, in the real 4-D world-through-time, color information of 4-D objects (e.g., cars and trees) is dominated by their 4-D spatiotemporal information, as properly stated by Adams *et al.* [26]. Second, the same consideration holds for a (2-D) image representation of the (4-D) world-through-time, where 2-D spatial (contextual) information dominates color information. To cope with the dominant 2-D spatial information in a (2-D) image, the human visual system employs modular arrays of multiscale 2-D local filters capable of providing a topology-preserving mapping of a (2-D) image [67]–[71], [74].

3) When Does Vision Go Symbolic? Inference Mechanisms in Human Vision: In the literature of psychophysics, according to Vecera and Farah, *preattentive image segmentation is an interactive (hybrid) inference process "in which top-down knowledge (e.g., familiarity) partly guides lower level processing"* ([75]; p. 1294). That is to say, *human vision is a symbolic hybrid (combined deductive and inductive) inference system where (symbolic) prior knowledge is injected into the sensory data interpretation process starting from the preattentive vision first stage* [18], [19].

In the computer vision literature, according to Marr "*(human) vision goes symbolic almost immediately, right at the level of (second-order derivative's) zero-crossing (raw primal sketch)... without loss of information*" ([5]; p. 343), which is consistent with the aforementioned quote by Vecera and Farah [75]. Unfortunately, in [5], the computer vision system proposed by Marr is unable to satisfy either one of the two aforementioned vision system requirements inspired by human vision. In particular, the Marr preattentive vision first stage is subsymbolic. It is split into a subsymbolic raw primal sketch and a subsymbolic full primal sketch, where: (I) the raw primal sketch consists of a hierarchy of subsymbolic primitives, namely, multiscale zero-crossings ([5]; pp. 54–59), followed by zero-crossing segments ([5]; p. 60) and level 1 image-tokens, comprising blobs (closed contours), edges, bars and discontinuities (terminations) ([5]; pp. 70–73), and (II) a full primal sketch, equivalent to perceptual grouping [75]–[77], where level 2 boundaries (e.g., texture boundaries) are detected between groups of tokens ([5]; pp. 53, 91–95). Marr never provided implementation details of his proposed subsymbolic raw primal sketch or subsymbolic full primal sketch. This apparent contradiction between Marr's computer vision system design (computational theory) specifications and his own implementation is not at all surprising. It accounts in general for the customary distinction between a model and the algorithm used to identify it [18].

4) Possible Relationships Between a Human Vision System and the ATCOR-SPECL and SIAM Prior Knowledge-Based Preclassifiers: Possible relationships between a human vision system, as it is described in Sections II-C1–II-C3, and the ATCOR-SPECL and SIAM prior knowledge-based preclassifiers, to be investigated in the Part 2 of this paper as alternative implementations of a preattentive vision first stage in a hybrid RS-IUS architecture [20], are highlighted as follows.

1) At the abstraction level of computational theory (system design), the hybrid RS-IUS architecture proposed in this paper is consistent with a human vision system conceived as a symbolic hybrid inference system where symbolic prior knowledge is injected right at the preattentive vision first stage (see Section II-C3).

2) In (2-D) images of the (4-D) world-through-time, 2-D spatial (contextual) information dominates color information (see Section II-C2). In traditional pixel-based RS-IUSs, the input data set is a 1-D sequence of pixel-specific data vectors where 2-D space (contextual) information is ignored. A pixel-based RS-IUS can perform accurately without 2-D spatial information in the image domain if and only if the image spatial resolution and time resolution are adequate to discriminate the target phenomenon under investigation based on (context-insensitive) color-through-time properties exclusively. It means that, to be considered useful, the application-independent ATCOR-SPECL and SIAM prior knowledge-based preclassifiers, which are pixel-based (context-insensitive) and eligible for use with any single-date RS imagery independent of its spatial resolution, must be considered as simple building blocks in a multistage RS-IUS architecture, i.e., they cannot be considered as standalone systems. In fact, their first-stage pixel-based (color-driven) preattentive image analysis must be followed by an attentive vision second stage, capable of (2-D) spatial analysis plus 1-D temporal

analysis of image data conditioned (driven, stratified) by first-stage spectral categories, equivalent to conventional color names to be community agreed upon [102], [103]. In terms of filling the information gap from sensory data to LC maps (refer to Section II-C1), the ATCOR-SPECL and SIAM prior knowledge-based preclassifiers map subsymbolic sensory data into semisymbolic spectral categories (refer to the further Section IV-B) based on single-date pixel-based MS (color) properties (spectral signatures) exclusively. The remaining information gap from semisymbolic spectral categories to LC classes must be filled by the RS-IUS' attentive vision second stage based on stratified spatiotemporal information.

We can conclude that, if compared with a human visual system, the degree of compatibility of the ATCOR-SPECL and SIAM prior knowledge-based preclassifiers, employed in support of the preattentive vision first stage of a hybrid RS-IUS architecture, is inferior to the degree of biological plausibility of an airplane compared to a bird. That said, from an engineering standpoint, the ATCOR-SPECL and SIAM deductive preclassifiers provide a realistic and feasible contribution to the development of automatic hierarchical RS-IUSs in operating mode, where a preattentional first-stage prior knowledge-based discretization of a continuous color space may be employed to better condition for numerical treatment an inherently difficult-to-solve second-stage attentive vision spatio-temporal analysis.

D. EO Big Data: Challenges and Opportunities

According to Section I, the secondary objective of this paper is to contribute to the development of a new generation of operational hybrid RS-IUSs capable of transforming large-scale multisensor multiresolution EO image databases into information products, in compliance with the QA4EO guidelines. The magnitude of EO data collected since the early 1970 s by a variety of spaceborne/airborne and *in situ* sensory data sources, at varying spatial extents and multiple spatial, temporal and spectral resolutions, is so phenomenal to be identified, by the present authors, as EO big data, in line with the terminology of IT.

In IT, the popular term "big data" identifies "a collection of data sets so large and complex that it becomes difficult to process using on-hand database management tools or traditional data processing applications. The challenges include capture, storage, search, sharing, analysis, and visualization" [78]. Among big data challenges, interpretation of observational data, i.e., the transformation of sensory data into information/knowledge products, has been historically investigated by both philosophical hermeneutics [21], [22] (refer to Section II-A) and psychophysical studies of perception [46] (refer to Section II-C).

According to the present authors, "big data" is a synonym of "central limit theorem." In statistics, the well-known central limit theorem states that [29], given certain conditions (typically random variables must be identically distributed), the sum (mean) of a sufficiently large number of independent random variables, each with a well-defined mean and well-defined variance (for example, one random variable is an LC class-specific distribution of pixel values in a RS image), tends to form a Gaussian distribution, where no "meaningful" or "natural" hidden data entities, clusters or (sub)structures can be identified [18], [19]. As a consequence of the central limit theorem, "big data" distributions are Gaussian-like, hence meaningful cluster/substructure detection in "big data" is inherently ill-conditioned in the Hadamard sense (refer to Section II-B). In other words, in "big data" sets, traditional inductive supervised or unsupervised data learning is extremely difficult or impossible to accomplish (refer to Section II-B).

These general considerations, driven from common knowledge in IT, may explain why, to date, EO big data assets are underemployed by the RS community. For example, the European Space Agency (ESA) estimates as 10% or less the percentage of RS images ever downloaded (which does not mean ever used) by stakeholders from its EO databases [18], [19]. It may mean that the RS discipline is still incapable of filling up the information gap from RS data to knowledge/information products (refer to Section II-C). To fill this information gap, data interpretation (cognitive) processes (related to the concept of equivocal "information-as-(an intepretation)process") dominate, i.e., are more difficult to solve than data transformation (e.g., data enhancement, data preprocessing) tasks (related to the concept of unequivocal "information-as-thing," refer to Section II-A). Typically, RS scientists and practitioners overlook their cognitive inadequacy to derive "operational, comprehensive, and timely knowledge/information products" from sensory data [1]–[3] by asking for more data of better quality, which actually makes their cognitive lack even worse. In practice, by overestimating its data interpretation capability the RS community is outpaced by the ever-increasing rate of collection of EO data of enhanced quality and quantity [10]–[19] (also refer to the further Section III).

To recapitulate, in agreement with common knowledge in IT, EO big data assets represent a huge opportunity/challenge for the RS interdisciplinary science. To be transformed into knowledge/information products in compliance with the QA4EO guidelines [1]–[3], EO big data require the development of a novel generation of hybrid inference systems in operating mode, capable of outperforming traditional inductive or deductive inference systems, whose limitations are well known (refer to Section II-B). As a realistic contribution to this challenge, this paper provides a theoretical and experimental assessment of the ATCOR-SPECL and SIAM prior knowledge-based preclassification software products in operating mode.

E. Probability and Nonprobability Sampling of a Geospatial Population

This paper requires that sample QIs, estimated from the ATCOR-SPECL and SIAM deductive preclassification maps, must be statistically valid (consistent), refer to Section I. By definition, an information map (e.g., a thematic map) is a reduced representation of a target geospatial population. To provide a statistically valid estimation of QIs from an information map representing a geospatial population [82], [83] (refer to

Section I), the following definitions of probability and nonprobability sampling protocol are required.

1) By definition, probability sampling must satisfy three necessary not sufficient conditions to deliver statistically valid sample estimates, i.e., sample estimates provided with the necessary probability foundation to permit generalization from the sample data set to the whole target geospatial population being sampled [82], [83]. 1) All inclusion probabilities must be greater than zero in the target geospatial population to be sampled. If some sampling units have an inclusion probability of zero, then the accuracy assessment does not represent the entire target region depicted in the map to be assessed and the results cannot be deemed statistically consistent. 2) The inclusion probabilities must be: a) knowable for nonsampled units and b) known for those units selected in the sample: since the inclusion probability determines the weight attached to each sampling unit in the accuracy estimation formulas, if the inclusion probabilities are unknown, so are the estimation weights. Probability sampling methods can be split into equal or variable (unequal) probability sampling methods. Unequal inclusion probabilities create no difficulties as long as they are known for sampled units and accounted for in the estimation formulas, but equal probability designs are advantageous in that they allow for simpler analysis. For example, an area sampling protocol selects polygons into the sample with an inclusion probability monotonically increasing with the polygon area [82], [83]. Noteworthy, no probability sampling is required to assess the degree of uncertainty in sample estimates [5].
2) Nonprobability sampling methods do not satisfy the requirements of probability sampling methods listed in this section above. According to the existing literature [82]: "unfortunately, examples of nonprobability sampling are common in accuracy assessment applications. Selecting reference locations by purposeful, convenient, or haphazard procedures does not allow the sampling design to determine the inclusion probabilities for each sampling unit. Such designs, therefore, are not probability samples. Purposefully, selecting training data for a supervised classification is a good example of a nonprobability sample. Such samples are acceptable for developing a land cover classification map, but often have limited use for accuracy assessment because the necessary probability foundation to permit generalization from the sample data to accuracy of the full population is lacking." To recapitulate, "it is possible to obtain useful information from nonprobability samples, but the limitations of such data should be recognized" [82]. For example, nonprobability sampling allows to assess the degree of uncertainty in sample estimates.
3) A protocol, defined as a sorted set of guidelines for good practice [3], encompasses a *structural knowledge* and a *procedural knowledge*, like in decision trees [55]. Structural knowledge is related to the content of the rule set while procedural knowledge is related to the order of presentation of rules. The definition of international protocols for best practices, such as the QA4EO guidelines [2], together with standardization, have been major challenges for the RS community [2], [3].

Unfortunately, in the RS literature there is a lack of probability sampling protocols adopted for the validation of RS data-derived products in compliance with the principles of statistics and the QA4EO guidelines. As a negative example of nonprobability sampling for map quality assessment not to be imitated, refer to [41].

A probability sampling protocol for thematic and spatial quality assessments of classification maps generated from EO images is proposed in [80] and adapted in Part 2 of this paper [20].

F. *QIO of an RS-IUS*

The test phase of a software product, which encompasses a QI selection stage, can be so relevant to absorb up to 50% of a project budget [93]. In this section, a possible list of mutually uncorrelated metrological/statistically-based QIOs is proposed and recommended for use by the Part 2 of this paper, to accomplish the experimental assessment and comparison of the ATCOR-SPECL and SIAM software products in operating mode [20].

Often forgotten in practice, the noninjective property of any metrological/statistically-based QI states that it is always possible to find two different instances of the same target phenomenon capable of generating the same QI value. For example, two different classification maps may provide the same map's overall accuracy value. This is tantamount to saying that no universal QI can exist [10], [19], which is in contrast with a significant segment of the existing literature, e.g., see [79] and [94]. Rather, a target-specific set of complementary statistically independent QIs must be selected and agreed upon by the scientific community.

To cope with EO big data challenges (refer to Section II-D), this paper provides an assessment of operational RS-IUSs in compliance with the principles of statistics, the QA4EO guidelines [2] and the GEO-CEOS land product accuracy validation criteria [3] (refer to Section I). These work requirements mean that the quality assessment of an RS-IUS should rely on a complete set of complementary metrological/statistically-based QIOs that are statistically independent, valid and significant. To be considered statistically significant, QIOs must be provided with a degree of uncertainty in measurement (refer to Section I). To be statistically valid (consistent), QIOs must be estimated from probability sampling of EO big data (refer to Section II-E).

Selected from the existing literature, a possible list of QIOs of an information processing system in operating mode is proposed as follows, to be community-agreed upon [10]–[19]. 1) Degree of automation (ease-of-use), monotonically decreasing with the number of system free-parameters to be user-defined based on heuristics. 2) Effectiveness, e.g., thematic accuracy and spatial accuracy of classification and segmentation maps generated from EO images [80]. 3) Efficiency, e.g., inversely related to computation time and memory occupation. 4) Robustness to changes in input parameters, if any free-parameter exists. 5) Robustness to changes in input data acquired across time, space and sensors. For example, refer to the CEOS land product accuracy validation stages 1–4 in [3], [4]. 6) Scalability, to cope with changes in input data specifications, sensors and user's requirements. 7) Timeliness, defined as the time between data acquisition and data-derived high-level product generation. For example, user interactions, such as those required to collect reference samples for training a supervised data learning system, increase timeliness [81]. 8) Costs, monotonically increasing with computer power and manpower.

To be termed operational, an information processing system must score high in every QIO of a set of community-agreed independent QIOs, e.g., refer to points 1) to 8) in the previous paragraph.

Unfortunately, experiments presented in large portions of the RS literature are affected by the following methodological drawbacks. 1) The sole mapping accuracy is selected from the possible set of mutually independent QIOs eligible for parameterizing RS-IUSs for assessment and comparison purposes. 2) Statistical estimates of the mapping accuracy are not provided with a degree of uncertainty in measurement, i.e., they have no statistical significance. 3) Statistical estimates of the mapping accuracy are not collected by means of a probability sampling strategy, hence they lack statistical consistency (refer to Section II-E). 4) Alternative RS data mapping solutions are tested exclusively in toy problems, defined in this paper as test data mapping problems featuring a small spatial scale (e.g., local scale) and/or a coarse semantic granularity, such that these test cases do not reflect the complexity of the existing "EO big data" archives (refer to Section II-D) that must be dealt with to comply with the QA4EO requirements [2] (refer to Section I). As a consequence of these experimental limitations, many RS-IUS implementations tested in the RS literature feature the following drawbacks. (I) A mapping accuracy which remains unknown in statistical terms and/or is unable to generalize from a sample data set to the whole target geospatial population being sampled. (II) A robustness to changes in the input data set which is unknown or appears questionable. (III) A scalability to real-world RS data applications at large (e.g., continental and global) spatial scale and fine semantic granularity which is unknown or appears questionable.

The conclusion of this section is that, in real-world RS data applications, different from toy problems at small spatial scale and/or coarse semantic granularity, published RS-IUSs are likely to score poorly in operating mode, because at least one of their OQI values is expected to score low.

III. Problem Recognition and Opportunity Identification: Compliance of Existing RS-IUS Commercial Software Products With the QA4EO Key Principles And Calibration/Validation (*Cal/Val*) Requirements

Adopted by the ongoing GEOSS implementation plan for years 2005–2015 [1], the international GEO-CEOS QA4EO recommendations promote the development of "operational, comprehensive, and timely knowledge/information products" from a variety of satellite, airborne, and *in situ* sensory data sources [2] (refer to Section I). To guarantee "the provision of and access to the Right Information, in the Right Format, at the Right Time, to the Right People, to Make the Right Decisions," the QA4EO guidelines require the successful implementation of two necessary and sufficient key principles [2]: (I) *Accessibility/Availability* and (II) *Suitability/Reliability* of RS data and data-derived knowledge/information products (refer to Section II-A). To accomplish these system requirements the GEO identified the need to develop a GEOSS data quality assurance strategy where calibration and validation (*Cal/Val*) activities become critical to data quality assurance and, thus, to data usability. According to the QA4EO guidelines [2], [3], the following *Cal/Val* activities are required.

1) *An appropriate coordinated program of calibration activities throughout all stages of a spaceborne mission, from sensor building to end-of-life, is considered mandatory to ensure the harmonization and interoperability of multisource multitemporal RS data* [2]. By definition, radiometric calibration is the transformation of dimensionless digital numbers (DNs) into a community-agreed physical unit of radiometric measure, e.g., TOA radiance (TOARD), TOA reflectance (TOARF), and spectral reflectance (SURF).
2) To satisfy validation requirements (e.g., accuracy validation [3]), *observational data and data-derived products generated in each step of a satellite-based information processing workflow must have associated with them a set of independent, quantifiable, metrological/statistically-based QIs, featuring a degree of uncertainty in measurement at a known degree of statistical significance*, to comply with the general principles of statistics and provide a documented traceability of the propagation of errors through the information processing chain in comparison with established "community-agreed reference standards" [2] (refer to Section II-F).

It is an indisputable fact that, to date, almost ten years from the launch of the GEOSS initiative, the RS community has been more successful in pursuing the first rather than the second GEOSS key principle. For example, in line with the GEOSS requirement of Accessibility/Availability of RS data and data-derived products, the U.S. 2008 free Landsat data policy has opened a new era of exploitation of the more than three million scenes stored in the U.S. Landsat archive [84]. On the other hand, the ever-increasing rate of collection of EO data of enhanced spatial, spectral and temporal quality outpaces the current ability of the RS discipline to transform EO big data assets into knowledge/information products (refer to Section II-D). This means that the GEOSS requirement of Suitability/Reliability of sensory data and data-derived products can still be considered far from being accomplished by the RS community.

To explain their different degrees of success, the first and second GEOSS key principles are analyzed at different levels of abstraction. At the abstraction level of knowledge/information representation, according to philosophical hermeneutics [21],

[22], the first GEOSS key issue is quantitative (unequivocal) and related to the Shannon concept of "information-as-thing" irrespective of its meaning [25]. As such, it is easier to deal with than the second GEOSS principle, which is qualitative (equivocal), since the latter has to deal with the meaning (interpretation, understanding) of sensory data and is related to the concept of "information-as-(an interpretation) process" [21], [22] (refer to Section II-A).

At the abstraction level of RS-IUS design, the second GEOSS key principle remains difficult to cope with also because *Cal/Val* activities are often neglected or ignored in the RS common practice. On theory, the RS community regards as common knowledge that *"the prerequisite for physically based, quantitative analysis of airborne and satellite sensor measurements in the optical domain is their calibration to spectral radiance"* ([95], p. 29). Moreover, according to related works [10]–[19], radiometric calibration is a necessary not sufficient condition for automatic interpretation of (for physical model-based inference from) EO imagery, refer to Section II-B. On the other hand, RS scientists, practitioners and institutions tend to overlook *Cal/Val* activities as necessary not sufficient preconditions for the harmonization of large-scale multitemporal multisensor EO datasets. For example, the European Commission Image 2000 product is a noncalibrated multisensor MS image mosaic at European scale, whose scientific usability for quantitative variable estimation is questionable or null [96]. To recover from this lack, the European Commission Image 2006 program includes radiometric calibration of multisensor MS images at European scale in its project requirements specification. However, in the Image 2006 project, no RS data-derived product validation policy is enforced [97].

To explain why radiometric calibration is neglected in the RS common practice, let us investigate the degree of compliance of RS-IUS commercial software products with the QA4EO key principles and *Cal/Val* requirements. Starting from the RS-IUS architectures proposed in Fig. 1, consider the: 1) two- or three-stage Trimble eCognition Developer [60], 2) one- or two-stage Pixel- and Segment-based versions of the Environment for Visualizing Images (ENVI) by ITT VIS [85], 3) one- or two-stage IDRISI Taiga, 4) one-stage ESRI ArcGIS, 5) ATCOR-2/3/4 [6]–[8], 6) one-stage PCI Geomatica (with an optional ATCOR for atmospheric correction), and 7) one- or two-stage ERDAS IMAGINE Objective (with an optional ATCOR for atmospheric correction). These commercial software packages for RS image processing/ understanding consist of large suites of options to choose from [18], [56]–[59]. Frequently considered overwhelming by nonexpert users, these large software suites allow selectable algorithms to be chosen, supervised, and combined by a user, based on heuristics, to form a user- and application-specific information processing workflow. Among these wide sets of selectable algorithms, several options may appear not particularly relevant, or be difficult to use (because they require lots of user interactions to run) or omit steps considered critical in a standard RS data processing chain (like those promoted by the QA4EO recommendations [2]). In practice, to favor flexibility considered necessary to develop customized solutions, these software suites promote an approach to RS image analysis closer to art, namely, empirical, qualitative and nonreproducible, than science, which is rigorous, quantitative and reproducible. For example, the large majority of selectable algorithms implemented in the RS-IUS commercial software products listed above, with the sole exception of the physical model-based ATCOR-2/3/4 toolbox [6]–[8], does not consider radiometric calibration as mandatory. This relaxed input data constraint means that, in these commercial software products, the large majority of selectable algorithms consist of statistical systems, hence the remaining small minority comprises physical models. Due to their inherent ill-posedness in the Hadamard sense [42], statistical systems are typically semiautomatic and site-specific [18], [45] (refer to Section II-B). Although statistical systems do not require as input observational data provided with a physical meaning, they may benefit from radiometric calibration in terms of robustness to changes in the input data set (refer to Section II-B). For example, in the ENVI commercial software toolbox [85], an atmospheric correction tool, called Fast Line-of-sight Atmospheric Analysis of Spectral Hypercubes (FLAASH), is available as an optional RS image preprocessing stage. As another example, in the PCI Geomatica and ERDAS RS data preprocessing workflows, a physical model-based ATCOR module can be optionally installed, etc.

The first conclusion about the RS-IUS commercial software products listed above is the following. In line with common knowledge in the machine learning community [24], since statistical model-based systems are inherently poorly-conditioned, semiautomatic and site-specific and require prior knowledge in addition to data to become better posed for numerical treatment (refer to Section II-B), then statistical systems available for selection in RS-IUS commercial software products, where they typically outnumber physical model-based options, are expected to be, *per se*, unable to cope with the well-known challenges of EO big data (refer to Section II-D). To become more successful, these statistical systems must be combined with physical models, to form hybrid inference systems capable of outperforming their individual components (refer to Section II-B). This consideration holds because at least one or more QIOs (e.g., timeliness, scalability, and robustness to changes in the input data set, refer to Section II-F) of any inductive data learning system, either supervised or unsupervised, whether or not it adopts an RS data radiometric calibration preprocessing stage in compliance with the QA4EO guidelines (refer to Section III), are expected to score low in real-world RS data mapping applications (refer to Section II-B), where EO big data assets (refer to Section II-D), different from unrealistic toy problems at small spatial scale and/or coarse semantic granularity (refer to Section II-F), are to be mapped.

In addition, RS-IUS commercial software products, such as those listed above, appear affected by a lack of selectable physical model-based inference systems, considered necessary to support, with prior knowledge in addition to data (in accordance with well-known principles of inductive inference, clearly stated by Mulier and Cherkassky [24], refer to Section II-B), the large majority of selectable options, consisting of statistical systems. This second conclusion about the RS-IUS commercial software products listed above is driven from the sole physical model found in this list, the ATCOR [6]–[8].

The core of the ATCOR consists of a radiative transfer model which is inverted to calculate as output directional surface reflectance (SURF) values starting from at-sensor (top-of-atmosphere, TOA) radiance (TOARD) values [9]. In the standard ATCOR implementation, the influence of surface type-specific bidirectional reflectance distribution function (BRDF) effects is not modeled. In the words of the ATCOR's authors [9], "ideally, an atmospheric and radiometric correction routine would result in BRDFs for all observed targets, as the BRDF is the unambiguous radiometric property of the Earth's surface. Unfortunately, imaging spectrometers rarely provide sufficient information to produce reliable BRDFs as most instruments acquire data for a single view geometry. Thus, a quantity not depending on the view geometry is of interest. The spectral albedo, i.e., the bihemispherical reflectance (BHR), is a value which is well suited for an unbiased view of the Earth's surface." In recent years, an "augmented" ATCOR implementation, sketched in Fig. 2, has been tested to retrieve spectral albedo in series with surface reflectance values starting from dimensionless DNs [9]. A peculiar aspect of this augmented ATCOR workflow, suitable for continuous variable estimation from RS data, is that categorical variables are generated as intermediate products by preliminary classification modules at several hierarchical stages (refer to Section II-A). In Fig. 2, data processing blocks identified as "preclassification" and "quantitative classification" are suitable for mapping semantic concepts from data, such as "clouds," "water," "vegetation," and "haze." Once estimated from sensory data, these categorical variables are further employed as input to processing modules capable of continuous (e.g., bio-physical) variable estimation (refer to Section II-B). That is to say, in the augmented ATCOR workflow shown in Fig. 2, the inherently poorly-conditioned inductive inference problem of continuous variable estimation from sensory data is accomplished on a symbolic stratified (driven-by-knowledge) basis to become better conditioned for numerical treatment (refer to Section II-B). In practice, the complete atmospheric correction and radiometric normalization scheme shown in Fig. 2 provides an additional source of experimental evidence supporting the recent conjecture, proposed in the RS literature [15], [80], that *categorical variables (e.g., LC and LCC maps) and continuous variables (e.g., spectral albedo, LAI and green biomass), conceived as two sides of the same coin, should be estimated from RS images alternately and iteratively, starting from a categorical variable estimation first stage* (refer to Section I). *Intuitively, MS image preclassification is preliminary to continuous variable estimation, which includes atmospheric correction, because the former task is "easier" to accomplish than the latter*. In fact, an expert photointepreter can successfully interpret (classify) an RS image irrespective of whether this image has been provided with a physical unit of radiometric measure through radiometric calibration. On the other hand, the RS literature clearly acknowledges that no spectral index (e.g., the normalized difference vegetation index, NDVI) should ever be computed as a quantitative proxy of a continuous biophysical variable (e.g., a LAI value), if no radiometric calibration has taken place, yet [45].

To summarize, capable of alternating categorical and continuous variable estimation from sensory data, the surface albedo

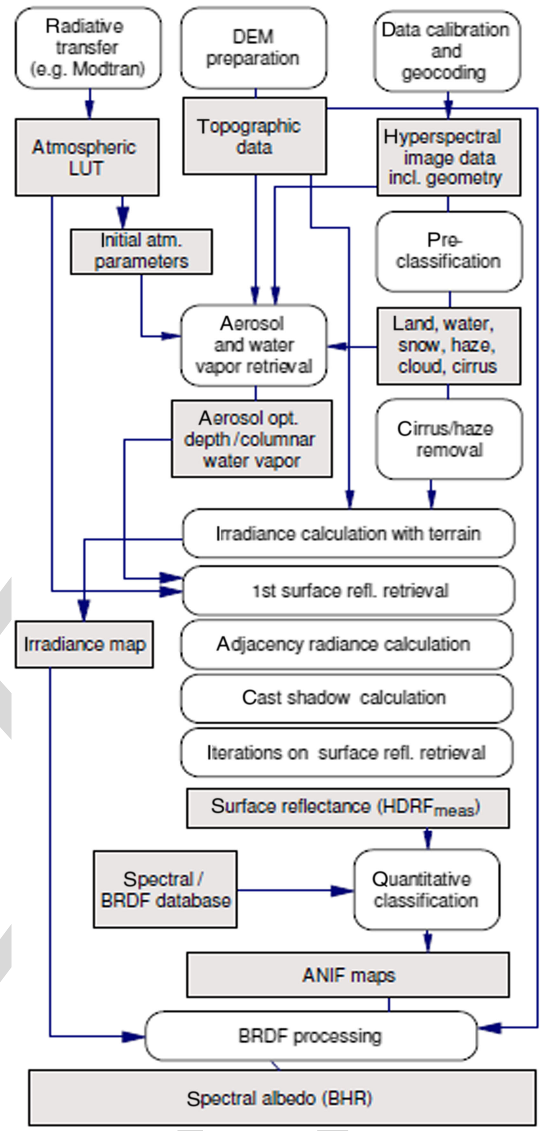

Fig. 2. A complete ("augmented") physical model-based system for RS data normalization combines a standard ATCOR workflow [6]–[9] with a novel bidirectional reflectance distribution function (BRDF) effect correction. Processing blocks are represented as circles and output products as rectangles. This workflow estimates categorical and continuous variables from sensory data alternately, starting from a prior knowledge-based pre-classification first stage, such as SPECL. Same as in [9], courtesy of Daniel Schläpfer, ReSe Applications Schläpfer.

estimation workflow shown in Fig. 2, based on an inverted radiative transfer model, is provided with a relevant degree of novelty in comparison with standard radiative transfer software products, like the Second Simulation of the Satellite Signal in the Solar Spectrum (6S) [86]. For example, in the 6S software tool, the land cover class-specific BRDF effects correction of RS imagery relies on ancillary thematic information, i.e., the 6S software product is per se unable to extract from the input RS image the surface types (e.g., ocean surface, vegetation and bare soil [86]) required as input to run the driven-by-knowledge BRDF correction phase.

This section concludes that, eligible for use as the physical model-based "preclassification" block in Fig. 2, *the ATCOR-SPECL and SIAM prior knowledge-based preclassifiers feature* *a wide application domain, encompassing not only categorical variable estimation from EO data* (as it is logical to expect from a preclassification system), *but also continuous variable estimation from EO data, in compliance with the Cal/Val activities considered mandatory by the QA4EO guidelines for both RS data preprocessing (data enhancement) and RS data processing (data understanding) phases* [2]. In other words, the ATCOR-SPECL and SIAM deductive preclassifiers appear as viable tools to accomplish not only automatic mapping of real-world EO big data sets (refer to Section II-D), in compliance with the QA4EO guidelines and the objectives of this paper (refer to Section I), but also RS image enhancement, as shown in Fig. 2. Existing examples of the SIAM applied to RS image preprocessing problems include stratified topographic correction [15], stratified atmospheric correction [6]–[8], stratified image mosaicking, stratified image co-registration, etc. [10]–[19] (refer to the further Section IV-A).

IV. Comparison of the ATCOR-SPECL and SIAM Software Products at the Four Levels of Understanding of an Information Processing System

Starting from the interdisciplinary nomenclature introduced in Section II, differences and similarities between the ATCOR-SPECL and SIAM software products can be investigated at the four levels of abstraction of an RS-IUS [5], [16], [18], [30], [87], namely: 1) computational theory (system architecture), 2) information/knowledge representation, 3) algorithms, and 4) implementation. Among these four levels of analysis, the first two are considered of fundamental importance for the success of any information processing system in operating mode (refer to Section I). In the words of Sonka *et al.*, "*the linchpin of success (of an information processing system) is addressing the (computational) theory (and information/knowledge representation [87]) rather than algorithms or implementation*" ([30]; p. 376).

A. Computational Theory

In Section I, the ATCOR-SPECL and SIAM software products are introduced as two alternative prior knowledge-based color space discretizers capable of providing a hybrid RS-IUS architecture with an injection of prior spectral knowledge, equivalent to color naming, right at the preattentive vision first stage, in compliance with human vision (refer to Section II-C). Common features of the two deductive image mapping systems are the following. 1) As physical models, they require as input a MS image provided with a physical unit of measure, namely, a MS image radiometrically calibrated into TOARF or SURF or surface albedo values (refer to Sections II-B and III). 2) They are context-insensitive, i.e., pixel-based, because color is the sole (0-D) pixel-specific information in a (2-D) image. All remaining visual properties are context-sensitive, e.g., texture [73], shape of image-polygons, and inter-object spatial relations [10]–[19], [46], [47], [61], [62]. 3) They are static, i.e., nonadaptive to input data, 4) one-pass, i.e., noniterative, 5) syntactic, i.e., rule-based [30], 6) semisymbolic, i.e., eligible for mapping a MS image into a discrete and finite set (legend) of spectral-based semiconcepts (refer to Section I), and 7) "fully automatic," because deductive inference systems require neither user-defined parameters nor training data sample to run [88] (refer to Section I).

Since they share the aforementioned list of system specifications, then the ATCOR-SPECL and SIAM systems can be used interchangeably in a hybrid RS-IUS workflow, such as those shown in Fig. 1(c) or 2. Although interchangeable, the ATCOR-SPECL and SIAM prior knowledge-based preclassifiers are not expected to perform the same, since their decision-tree design and implementation are completely different, in terms of both structural and procedural knowledge (refer to Section II-E).

A novel three-stage hybrid RS-IUS architecture, shown in Fig. 1(c), whose preattentive vision first stage employs a prior knowledge-based preclassifier provided with feedback loops [10]–[19], is described as follows.

1) An EO image preprocessing stage zero, suitable for MS image enhancement, including a mandatory MS image radiometric calibration of DNs into TOARF values, in compliance with the QA4EO guidelines. Although SURF values, considered as a special case of TOARF values in very clear sky conditions and flat terrain conditions [12], [80], [89], i.e., $\mathrm{TOARF} \supseteq \mathrm{SURF}$, such that $\mathrm{TOARF} \approx \mathrm{SURF} + \mathrm{atmospheric}$ "noise," are allowed as input, they are not mandatory, i.e., atmospheric correction is not considered a MS image preprocessing requirement.
2) A physical model-based symbolic context-insensitive (pixel-based) preattentive vision first stage, like the ATCOR-SPECL or the SIAM prior knowledge-based preclassifier. An injection of prior knowledge in the preattentive vision first stage makes the inherently poorly-conditioned EO image interpretation problem better posed for numerical treatment (refer to Section II-B), in agreement with the Marr intuition that vision goes symbolic right at the level of the raw primal sketch [5] (refer to Section II-C).
3) A second-stage battery of attentive vision context-sensitive stratified (driven-by-knowledge) application-, sensor- and LC/LCC class-specific feature extractors (e.g., multiscale texture is investigated exclusively in the image portion masked by the first-stage spectral category "vegetation," in order to split spectral type "vegetation" into two LC classes, namely, low-texture "grassland" and high-texture "forest" [61], [62]) and one-class LC/LCC classification modules (e.g., if a first-stage spectral category mask is "vegetation" and the second-stage "vegetation" masked data feature extractor is "high texture," then "forest").
4) A feedback mechanism between the preattentive vision first stage, the attentive vision second stage and the RS image preprocessing stage zero. Existing examples of these feedback loops are stratified topographic correction [15], stratified atmospheric correction [6]–[8], stratified image mosaicking, stratified image co-registration, and cloud/cloud-shadow masking [10]–[19].

This novel hybrid RS-IUS design [see Fig. 1(c)] is alternative to the two-stage hybrid RS-IUS architecture proposed by Shackelford and Davis [61], [62], whose first stage is a nonadaptive statistical classifier, namely, a plug-in parametric ML classifier (refer to Section II-B), and to state-of-the-art two-stage noniterative GEOBIA system [see Fig. 1(b)] and three-stage iterative GEOOIA system architectures [18], [19] (refer to Section II-B), where: 1) the preattentive vision first stage consists of an unlabeled data learning algorithm for image segmentation [32]–[34], [55]–[60], which is inherently poorly-posed [24] and is, therefore, semiautomatic and site-specific [45]; and 2) prior knowledge, if any, is injected exclusively at the attentive vision second stage, if and only if this second stage is implemented as a static image-object-based decision-tree classifier. If no prior knowledge is employed at the GEOBIA/GEOOIA attentive vision second stage, because it is implemented as an inductive data learning classifier (e.g., an artificial neural network classifier, a support vector machine classifier [41], a nearest-neighbor classifier, an adaptive decision-tree classifier, and a radial basis function network for classification [24], [29]), then the GEOBIA/GEOOIA system implementation is fully inductive at both first and second stages, which means that the GEOBIA/GEOOIA system, due to its inherent ill-posedness, is semiautomatic and site-specific in common practice (refer to Section II-B). This line of reasoning justifies the low productivity of many GEOBIA/GEOOIA systems increasingly observed in the existing literature [56], [57], which makes them inadequate to cope with large-scale RS image databases.

B. *Information/Knowledge Representation*

The ATCOR-SPECL and SIAM software products are compared in terms of: 1) input MS data requirements and 2) output preclassification map's legend.

1) Input MS Data Requirements Specification: The physical model-based ATCOR-SPECL and SIAM prior knowledge-based preclassifiers require as input MS images radiometrically calibrated into a physical unit of radiometric measure (refer to Section II-B), in compliance with the *Cal/Val* requirements of the QA4EO guidelines [2] (refer to Section III). In particular, SIAM requires as input a MS image radiometrically calibrated into TOARF or SURF or surface albedo values, where SURF is a special case of TOARF in very clear sky conditions and flat terrain conditions [12], [80], [89], i.e., $\mathrm{TOARF} \supseteq \mathrm{SURF}$, such that $\mathrm{TOARF} \approx \mathrm{SURF} + \mathrm{atmospheric}$ "noise." It means that an LC class-specific family of spectral signatures in TOARF values forms a buffer area (envelope) which includes, as a special case, the family of "ideal" (atmospheric noiseless) spectral signatures in SURF values for that same LC class, see Fig. 3.

In practice, SIAM is capable of recognizing surface types in RS images by "looking through" atmospheric effects, like the presence of haze and thin clouds [10]–[19]. This "look-through" capability is due to the fact that the original spectral prior knowledge base of the SIAM consists of a reference dictionary of spectral signatures in TOARF values, where relation $\mathrm{TOARF} \approx (\mathrm{SURF} + \mathrm{atmosphericnoise})$ holds, whereas traditional libraries of spectral signatures are in SURF values (measured at the ground level) exclusively, i.e., they are atmospheric noise-free. Well-known examples of reference dictionaries of spectral signatures in (atmospheric noise-free) SURF values, such as the U.S. Geological Survey (USGS) mineral and vegetation spectral libraries, the Johns Hopkins University spectral library and the Jet Propulsion Laboratory mineral spectral library [6]–[9], can be found in the existing literature, e.g., refer to [90] (p. 273) or in commercial software products [85]. Being provided with an (implicit) atmospheric noise model, the SIAM is expected to be robust to the presence of atmospheric effects. This means that SIAM does not consider preliminary atmospheric correction as mandatory because SIAM is knowledgeable on how to cope with RS data affected by atmospheric noise.

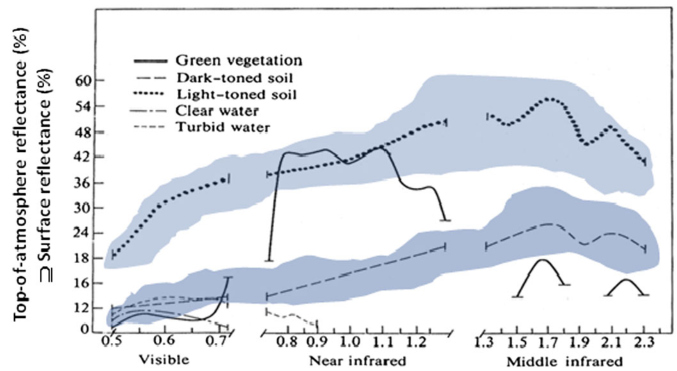

Fig. 3. Land cover (LC)-class specific families of spectral signatures in TOA reflectance (TOARF) values form buffer areas (envelopes) which include surface reflectance (SURF) values as a special case in clear sky and flat terrain conditions.

Unlike the SIAM reference dictionary of spectral signatures in TOARF values, the ATCOR-SPECL rule set has been developed starting from a prior knowledge base of reference spectral signatures in SURF values [6], [91], which means that the ATCOR-SPECL requires atmospheric correction as a mandatory preprocessing stage. In general, atmospheric correction is inherently poorly-conditioned and, therefore, difficult to solve. In practice, atmospheric correction requires user-supervision to become better posed for numerical treatment, also refer to Fig. 2 [6]–[9]. Although it requires SURF values as input data, the ATCOR-SPECL software product is expected to be able to cope with (to look-through) input images in TOARF values, when atmospheric effects are those typical of clear or very clear sky conditions and topographic effects are negligible, such that $\mathrm{TOARF} \approx \mathrm{SURF}$ [89].

2) First-Stage Output Semisymbolic Information Primitives: In a community-agreed ontology of the 4-D world-through-time (refer to Section II-C), e.g., in an LC or LCC map's legend (vocabulary), each ontological concept, e.g., each LC or LCC class name in the vocabulary, identifies a specific class of surface objects in the 4-D world-through-time featuring specific 4-D spatio-temporal properties, together with spectral (color) properties. In general, *LC class-specific spatio-temporal information dominates color information* [26] (refer to Section I), which is the reason why achromatic vision can be very successful despite the absence of color information.

In a preclassification map generated by the ATCOR-SPECL and SIAM software products from a single-date MS imagery, the map legend consists of a discrete and finite set of semisymbolic informational primitives, called color names, color-based inference categories, spectral-based semiconcepts, spectral categories or spectral endmembers, such as "vegetation," "bare soil or built-up," and "water or shadow" [10]–[19], [26]. Each spectral-based semiconcept can be mapped onto (matched with) one or more LC classes whose spectral properties can overlap, irrespective of spatio-temporal properties capable of disambiguating these LC classes (refer to Section I). In other words, spectral-based semiconcepts are single-date and pixel-specific, i.e., they ignore the (dominant) 4-D spatio-temporal information carried by LC classes, but exclusively investigate the (dominated) color properties of LC classes. As a consequence, the semantic meaning of a spectral-based semiconcept (e.g., "vegetation") is: 1) superior to zero, where zero is the semantic information conveyed by subsymbolic image features, i.e., image-objects (image-polygons) or, vice versa, image-contours (since image contour detection is the dual task of image segmentation and they are both poorly-posed [10]–[19]); and 2) equal or inferior to the semantic meaning of concepts in the attentive vision second stage, i.e., LC classes, e.g., "needle-leaf forest," belonging to a world model, namely, a spatio-temporal ontology of the 4-D world-through-time.

Hence, in general, one spectral-based semiconcept can be associated with none, one or many LC classes (refer to Section I). For example, spectral category "strong vegetation" can be linked to LC classes "grassland" or "agricultural field" or "forest," just like "*endmember fractions cannot always be inverted to unique class names*" ([26], p. 147). Analogously, one LC class can encompass different color discretization levels, e.g., the LC class "deciduous forest" can look like several tones of green equivalent to the SIAM's color quantization levels (spectral categories, color names) "strong vegetation," "average vegetation," and "dark vegetation." This means that, in general, a finite set of many-to-many associations holds between spectral-based semiconcepts in the (2-D) image domain and the reference LC classes belonging to a spatio-temporal ontology of the 4-D world-through-time [80]. Special cases of many-to-many inter-vocabulary relations are one-to-many, many-to-one and one-to-one relations. Many-to-many inter-legend relations convey mapping information because only all-to-all inter-legend "correct" entries do not (like if every spectral category were mapped onto all LC classes). For example, proposed in [80], an original Categorical Variable Pair Similarity Index (CVPSI) provides an estimated value, around 50%, of the degree of match between the SIAM's vocabulary and the LC class legend adopted by the USGS 2006 National Land Cover Data map, also refer to Fig. 1(c).

At a finer level of detail, SIAM delivers as output preclassification maps at various levels of color discretization, namely, fine, intermediate and coarse, where prior knowledge-based color quantization levels depend on the spectral resolution of the imaging sensor. At coarse granularity, SIAM's spectral categories belong to the following six parent spectral categories (also called super-categories) or major spectral endmembers: 1) "Clouds," 2) "Either snow or ice," 3) "Either water or shadow," 4) "Vegetation," equivalent to "either woody vegetation or cropland or grassland (herbaceous vegetation) or (shrub and brush) rangeland," 5) "Either bare soil or built-up," and 6) "Outliers."

These SIAM super-categories can be compared with the four reference endmembers, namely, "green vegetation," "nonphotosynthetic vegetation" (e.g., woody material on the ground together with dead or dying leaves), "soil," and "shadow," derived from laboratory surface reflectance spectra by Adams *et al.* in spectral mixture analysis [26].

Due to the presence of class "Outliers" ("Unknowns"), SIAM provides a mutually exclusive and totally exhaustive mapping of the input MS image into a discrete and finite vocabulary (legend) of color names, in line with the Congalton and Green requirements of a classification scheme [92]. It is noteworthy that, although the definition of a rejection rate is a well-known objective of any RS image classification system, e.g., refer to [26] and [90], RS image classifiers are often applied without any outlier detection strategy.

Similar considerations hold for the ATCOR-SPECL preclassifier, refer to the ATCOR-SPECL legend shown in Table I. For example, to identify information primitives of an ATCOR-SPECL's output map, the most recent ATCOR user guides, like [7] and [8], adopt the same term, "spectral categories," originally proposed in the SIAM literature to differentiate spectral-based semiconcepts from traditional LC classes [10]–[19]. According to [6]–[8], revised by Richter [91], the ATCOR-SPECL static decision-tree preclassifier consists of a sorted set of 19 spectral categories, including class "unknowns" (refer to Table I), in compliance with the Congalton and Green requirements of a classification scheme [92].

C. Algorithm Design

In [93], algorithm design is defined as "everything, but code." This definition is recalled to point out that, although they belong to the same family of spectral knowledge-based preclassifiers (refer to Section IV-A), capable of transforming subsymbolic observational data into semisymbolic spectral categories (refer to Section IV-B), the ATCOR-SPECL and SIAM software products are totally different in terms of decision-tree design, comprising both structural and procedural knowledge (refer to Section II-E), irrespective of implementation.

Sonka *et al.* describe aspects of image-object labeling through artificial intelligence in terms of syntactic pattern recognition ([30]; p. 285). In syntactic pattern recognition, the following considerations hold.

1) Elementary properties of the syntactically described objects from a given class are called primitives. Relations between objects may be modeled as hierarchical relational structures.
2) A class-specific description language is the set of all words that may be used to describe objects from one class, based on information primitives. For example, in written language, words of the language are constructed from letters and the set of all letters is called the alphabet. Letters are equivalent to information primitives and the words of the language are created from a collection of the alphabet's letters.

TABLE I
SPECTRAL RULES AND PSEUDO-COLORS OF THE LEGEND ADOPTED BY THE ATCOR-SPECL PRIOR KNOWLEDGE-BASED PRECLASSIFIER [6], [91]

Index	**Spectral categories**	**Spectral rule** (based on reflectance measured at Landsat TM central wave bands: b1 is located at 0.48 µm, b2 at 0.56 µm, b3 at 0.66 µm, b4 at 0.83 µm, b5 at 1.6 µm, and b7 at 2.2 µm)	**Pseudo-color**
1	Snow/ice	$b4/b3 \leq 1.3$ AND $b3 \geq 0.2$ AND $b5 \leq 0.12$	
2	Cloud	$b4 \geq 0.25$ AND $0.85 \leq b1/b4 \leq 1.15$ AND $b4/b5 \geq 0.9$ AND $b5 \geq 0.2$	
3	Bright bare soil/ sand/cloud	$b4 \geq 0.15$ AND $1.3 \leq b4/b3 \leq 3.0$	
4	Dark bare soil	$b4 \geq 0.15$ AND $1.3 \leq b4/b3 \leq 3.0$ AND $b2 \leq 0.10$	
5	Average vegetation	$b4/b3 \geq 3.0$ AND ($b2/b3 \geq 0.8$ OR $b3 \leq 0.15$) AND $0.28 \leq b4 \leq 0.45$	
6	Bright vegetation	$b4/b3 \geq 3.0$ AND ($b2/b3 \geq 0.8$ OR $b3 \leq 0.15$) AND $b4 \geq 0.45$	
7	Dark vegetation	$b4/b3 \geq 3.0$ AND ($b2/b3 \geq 0.8$ OR $b3 \leq 0.15$) AND $b3 \leq 0.08$ AND $b4 \leq 0.28$	
8	Yellow vegetation	$b4/b3 \geq 2.0$ AND b2 ≥_b3 AND $b3 \geq 8.0$ AND $b4/b5 \geq 1.5$[a]	
9	Mix of vegetation/ soil	$2.0 \leq b4/b3 \leq 3.0$ AND $0.05 \leq b3 \leq 0.15$ AND $b4 \geq 0.15$	
10	Asphalt/dark sand	$b4/b3 \leq 1.6$ AND $0.05 \leq b3 \leq 0.20$ AND $0.05 \leq b4 \leq 0.20$[a] AND $0.05 \leq b5 \leq 0.25$ AND $b5/b4 \geq 0.7$[a]	
11	Sand/bare soil/ cloud	$b4/b3 \leq 2.0$ AND $b4 \geq 0.15$ AND $b5 \geq 0.15$[a]	
12	Bright sand/bare soil/cloud	$b4/b3 \leq 2.0$ AND $b4 \geq 0.15$ AND ($b4 \geq 0.25$[b] OR $b5 \geq 0.30$[b])	
13	Dry vegetation/ soil	($1.7 \leq b4/b3 \leq 2.0$ AND $b4 \geq 0.25$[c]) OR ($1.4 \leq b4/b3 \leq 2.0$ AND $b7/b5 \leq 0.83$[c])	
14	Sparse veg./soil	($1.4 \leq b4/b3 \leq 1.7$ AND $b4 \geq 0.25$[c]) OR ($1.4 \leq b4/b3 \leq 2.0$ AND $b7/b5 \leq 0.83$ AND $b5/b4 \geq 1.2$[c])	
15	Turbid water	$b4 \leq 0.11$ AND $b5 \leq 0.05$[a]	
16	Clear water	$b4 \leq 0.02$ AND $b5 \leq 0.02$[a]	
17	Clear water over sand	$b3 \geq 0.02$ AND $b3 \geq b4 + 0.005$ AND $b5 \leq 0.02$[a]	
18	Shadow		
19	Not classified (outliers)		

[a] These expressions are optional and only used if b5 is present.
[b] Decision rule depends on presence of b5.
[c] Decision rule depends on presence of b7 [8].

3) A class-specific description grammar is the set of (substitution) rules that must be followed when words of the class-specific description language are constructed from letters. In other terms, each class consists only of objects whose syntactic description is syntactically correct according to the particular class grammar. In the written language example, although many words may be used together, only those which follow the correct grammar will form a coherent sentence.
4) Syntactic recognition is a process that looks for the class-specific grammar that can generate the syntactic word or phrase which describes an unknown object.
5) (Qualitative) syntactic object description should be used whenever (quantitative) statistical feature description is not able to represent the complexity of the target objects and/or when there are inter-object relations, like *part-of* or *subset-of*, difficult to learn from data by means of inductive data learning algorithms and that typically require significant human interaction to be identified.

In the aforementioned terminology of syntactic pattern recognition systems, both the ATCOR-SPECL and SIAM deductive decision-tree preclassifiers are built upon a physical knowledge base of families (envelops) of real-world spectral signatures per surface type (e.g., "bare soil or built-up"), so that a sorted set of land surface type-specific grammars (hierarchical decision-tree) is constructed.

In the SIAM software product, a spectral category-specific grammar is a combination of two information primitives capable of describing the family of spectral signatures belonging to that surface type (see [11] for full details). The first spectral primitive is the so-called "spectral rule" whose aim is to describe the shape of a buffer zone (envelope) of a surface type-specific family of spectral signatures in TOARF values, irrespective of intensity (see Fig. 2). In particular, a spectral rule defines a buffer zone of spectral tolerance, irrespective of the absolute intensity of spectral bands, by means of relational operators ($<$, $>$, $\leq$, $\geq$) between spectral bands. The second spectral primitive is a spectral fuzzy set (e.g., low, medium, and high) extracted from the intensity of scalar spectral variables, namely, spectral bands or spectral indexes. To recapitulate, a surface type-specific grammar is a combination of logical operators (AND, OR, NOT) with one or more spectral rules and/or one or more spectral fuzzy sets, capable of modeling the shape and the radiometric intensity of the surface type-specific MS envelope of spectral signatures [11].

Unlike SIAM, where a spectral category-specific grammar consists of a logical (AND, OR, NOT) combination of one or more spectral rules and spectral fuzzy sets [11], each ATCOR-SPECL's category-specific grammar consists of a single spectral rule per spectral category [6]–[8], see Table I.

Since the rule complexity of the SIAM expert system is superior to that of the ATCOR-SPECL, the former is expected to be more accurate than the latter at the cost of a higher implementation complexity and computation time.

To conclude this section, let us point out the algorithmic difference between the ATCOR-SPECL and SIAM prior knowledge-based preclassifiers and the popular spectral mixture analysis for MS image classification [26]. In spectral unmixing, the so-called (endmember) fraction categories are detected by category-specific boundaries established sequentially and in a particular order by an application developer in an E-dimensional measurement space, where E is the total number of reference endmembers, such that E is always less or equal than the number of spectral bands minus 1. For example, in the work of Adams *et al.* [26], dealing with 7-band Landsat images, the free number of spectral endmembers E is set equal to four, to allow the endmember space be rotated by the application developer on the computer screen to show any desired projection. On the contrary, the prior knowledge-based preclassification decision trees implemented in the ATCOR-SPECL and SIAM software products consist of dozens of prior knowledge-based category-specific grammars, whose inputs are spectral bands and spectral indexes, but never reference endmembers. Rather, the ATCOR-SPECL and SIAM expert systems, consisting of prior knowledge-based color discretization levels equivalent to data- and application-independent spectral endmembers, are suitable for automatic preclassification of hyperspectral images as a viable deductive alternative to state-of-the-art inductive algorithms for spectral endmember learning from hyperspectral data [104].

D. Implementation

The two ATCOR-SPECL and SIAM deductive decision-tree preclassifiers are totally different at the abstraction level of algorithm design (refer to Section IV-C), encompassing the list of category-specific grammars (structural knowledge [55]) and their order of presentation (procedural knowledge [55]). As a consequence, they are completely different at the implementation level of analysis.

According to [6]–[8], revised by Richter [91], the static decision-tree preclassifier currently implemented in the ATCOR-SPECL secondary software product consists of a sorted set of 19 spectral category-specific grammars (refer to Table I) which includes class "unknowns" (refer to Section IV-B2). In terms of semantic granularity the ATCOR-SPECL is coarser than the SIAM (vice versa, the semantic cardinality of the former is inferior to that of the latter), which means that the implementation complexity of the latter's decision tree is greater than that of the former (also refer to Section IV-C).

To the best of these authors' knowledge, *the SIAM software product is the first semisymbolic expert system* (refer to Section II-B)*, made available to the RS community for operational use in a RS-IUS preattentive vision first stage* (refer to Section II-C)*, capable of accomplishing multiscale image segmentation and multigranule image preclassification simultaneously, automatically and in near real-time* [10]–[19]. The extraction of a (subsymbolic) image segmentation map (where subsymbolic image-objects are identified as, say, segment 1, segment 2, etc.) from a binary or multilevel image (e.g., a thematic map) can be accomplished by a traditional well-posed (deterministic) automatic (requiring no user interaction) two-pass connected-component image labeling algorithm, e.g., refer to [30] (p. 197). In practice, a unique (subsymbolic) segmentation map can be generated from a multilevel image, like a thematic map, but the contrary does not hold, e.g., different thematic maps can generate the same segmentation map, i.e., no unequivocal thematic map can be inferred from a segmentation map [18], [19]. In other words, a realistic alternative to the (e.g., eCognition's) generation of an inherently poorly-conditioned, semiautomatic and site-specific multiscale segmentation map from an input subsymbolic MS image is the automatic well-posed generation of a multiscale segmentation map from a multilevel semisymbolic preclassification map, featuring several degrees of color discretization (e.g., fine, intermediate and coarse), which has been automatically generated by a prior knowledge-based multigranule preclassifier from an input MS image.

SIAM is implemented as an integrated system of six subsystems, including one "master" Landsat-like subsystem plus five "slave" (down-scale) subsystems, whose spectral resolution overlaps with Landsat's, but is inferior to Landsat's, refer to Table II. Noteworthy, the expression "Landsat-like MS image" adopted in this paper means: "an MS image whose spectral resolution mimics the spectral domain of the 7 bands of the Landsat family of imaging sensors," i.e., a spectral resolution where bands visible blue (B), visible green (G), visible red (R), near infra-red (NIR), medium infra-red 1 (MIR1), medium infra-red 2 (MIR2) and thermal infra-red (TIR) overlap (which does not mean coincide) with Landsat's.

The aforementioned SIAM's six subsystems are summarized in Table II. The output spectral categories detected at the fine, intermediate and coarse color discretization levels by the SIAM's six subsystems, described in Table II, are summarized in Table III.

With regard to the SIAM implementation, in [11] enough information is provided for the crisp L-SIAM implementation

T2:1 TABLE II
T2:2 LIST OF SPACEBORNE/AIRBORNE SENSORS ELIGIBLE FOR USE WITH THE SIAM SYSTEM OF SYSTEMS

SIAM system of systems		**B—(E)TM1, 0.45–0.52 (µm)**	**G—(E)TM2, 0.52–0.60 (µm)**	**R—(E)TM3, 0.63–0.69 (µm)**	**NIR—(E)TM4, 0.76–0.90 (µm)**	**MIR1—(E)TM5, 1.55–1.75 (µm)**	**MIR2—(E)TM7, 2.08–2.35 (µm)**	**TIR—(E)TM6, 10.4–12.5 (µm)**	**SR (m)**	**Rad. Cal. Y/N, C/I**	**Pan SR (m)**	**Notes**
L-SIAM, Input bands: 7 – B, G, R, NIR, MIR1, MIR2, and TIR. Output Sp. Cat.: 96/48/18	**Landsat-4/-5 TM**	×	×	×	×	×	×	×	30	Y-C		Refer to Table I in [11]
	Landsat-7 ETM+	×	×	×	×	×	×	×	30	Y-C	15	Same as above.
	Landsat-8 OLI+TIRS	×	×	×	×	×	×	×	30	Y-C	15	
	MODIS	×	×	×	×	×	×	×	250, 500, 1000	Y-C		Same as above.
	ASTER		×	×	×	×	×	×	15-30	Y-C		Same as above.
	CBERS-2B	×	×	×	×	×	×	×		N		
	APEX	×	×	×	×	×	×		1.8	Y		Airborne hyperspectral, 285 bands
	AVIRIS	×	×	×	×	×	×		e.g., 20	Y-?		Airborne hyperspectral, 224 bands, managed by Jet Propulsion Laboratory (JPL)
	MIVIS	×	×	×	×	×	×	×	e.g., 1.64	Y-?		Airborne hyperspectral, 102 bands, managed by CNR, Italy
	Sentinel-2 MSI	×	×	×	×	×	×		10 (B, G, R, NIR), 20 (MIR 1, MIR2)	?		13 bands, from VIS to MIR. To be launched in 2015?
	Sentinel-3 SLSTR		×	×	×	×	×	×	500	?		9 bands, from VIS to TIR + 2 (active fire). To be launched in 2015?
	WorldView-3	×	×	×	×	×	×		MS: 1.24, SWIR: 3.7	Y-C	0.3	16 bands, from VIS to SWIR. Launched in Aug. 2014.
S-SIAM, Input bands: 4—G, R, NIR, MIR1. Output Sp. Cat.: 68/40/15	**SPOT-4 HRVIR**		×	×	×	×			20	Y-I	10	Refer to Table II in [11].
	SPOT-5 HRG		×	×	×	×			10	Y-I	2.5–5	Same as above.
	SPOT-4/-5 VMI		×	×	×	×			1100	Y-I		Same as above.
	IRS-1C/-1D LISS-III		×	×	×	×			23.5	Y-I		
	IRS-P6 LISS-III		×	×	×	×			23.5	Y-I		
	IRS-P6 AWiFS		×	×	×	×			56	Y-I		

TABLE II
CONTINUED

AV-SIAM, Input bands: 4 —R, NIR, MIR1, TIR. Output Sp. Cat.: 83/43/17	**NOAA AVHRR**			×	×	×		×	1100	Y		Refer to Table II in [11].
	MSG			×	×	×		×	3000	Y		Same as above.
	NASA-NOAA NPP VIIRS			×	×	×	×	×	375	Y-C		
AA-SIAM, Input bands: 5 —G, R, NIR, MIR1, TIR. Output Sp. Cat.: 83/43/17	**ENVISAT AATSR**		×	×	×	×		×	1000	Y		Same as above.
	ERS-2 ATSR-2		×	×	×	×		×	1000	Y		
Q-SIAM, Input bands: 4 —B, G, R, NIR. Output Sp. Cat.: 61/28/12	**IKONOS-2**	×	×	×	×				4	Y-C	1	
	QuickBird-2	×	×	×	×				2.4	Y-C	0.61	
	GeoEye-1	×	×	×	×				1.64	Y	0.41	
	OrbView-3	×	×	×	×				4	N	1	
	SPOT-6/7	×	×	×	×				6	Y-I	1.5	
	Pleiades-1A/1B	×	×	×	×				2	Y-I	0.5	
	RapidEye-1 to -5	×	×	×	×				6.5	Y-I		
	ALOS AVNIR-2	×	×	×	×				10	Y-C		
	KOMPSAT-2	×	×	×	×				4	N	1	
	TopSat	×	×	×	×				5	N	2.5	
	FORMOSAT -2	×	×	×	×				8	Y-?	2	
	Huan Jing satellite constellation, HJ-1A / HJ-1B, payload: WVC.	×	×	×	×				30	Y-C		Wide View CCD cameras (WVC).
	ENVISAT MERIS	×	×	×	×				300	Y-?		Super-spectral, 15 bands
	Sentinel-3 OLCI	×	×	×	×				300, 1200			Super-spectral, 21 bands. To be launched in 2015?
	Leica ADS-40/80	×	×	×	×				0.25	Y-?	0.25	Airborne, 4 bands + PAN
D-SIAM, Input bands: 3 —G, R, NIR. Output Sp. Cat.: 61/28/12	**Landsat-1/-2/-3/-4/-5 MSS**		×	×	×				79	Y-C		
	IRS-P6 LISS-IV		×	×	×				5.8	Y-I		
	SPOT-1/-2/-3 HRV		×	×	×				20	Y-I	10	
	DMC		×	×	×				22-32	Y-C		

Acronyms: Y, Yes; N, No; C, Complete; I, Incomplete (radiometric calibration offset parameters are set to zero); (E)TM, (Enhanced) Thematic Mapperl; B, Blue; G, Green; R, Red; NIR, Near Infra-Red; MIR, Medium Infra-Red; TIR, Thermal Infra-Red; SR, Spatial Resolution; and Pan, Panchromatic.
Adopted acronyms: SPOT, Satellite Pour l'Observation de la Terre; NOAA, National Oceanic and Atmospheric Administration (NOAA); AVHRR, Advanced Very High Resolution Radiometer; AATSR, ENVISAT Advanced Along-Track Scanning Radiometer; Q, QuickBird; DMC, Disaster Monitoring Constellation.
Column highlight color: Blue columns are related to visible channels typical of water and haze; Green column identify the NIR band, typical of vegetation; Brown columns are related to MIR channels, characteristic of bare soils; and Red column: TIR channel, useful to detect fire.

to be reproduced. The down-scale S-SIAM, AV-SIAM and Q-SIAM versions, generated from the "master" L-SIAM implementation (refer to Table II), are described in [12]–[14]. In [17], the crisp-to-fuzzy SIAM transformation is explained in detail. It is noteworthy that since its first 2006 release presented in [11], L-SIAM has increased its number of output spectral categories from 46 to 96 (see Table II). This progressive, but slow, increase in the number of spectral categories detected by the sequence of "master" L-SIAM implementations proposed to the RS literature in recent years shows that, in line with theory [45], [55] (refer to Section II-B), there is a slow "learning curve" in the development and fine-tuning of physical models, such as the ATCOR-SPECL and SIAM prior knowledge-based preclassifiers.

TABLE III
SIAM SYSTEM OF SIX SUBSYSTEMS

<table>
<tr><th rowspan="2">SIAM</th><th rowspan="2">Input bands (B: Blue, G: Green, R: Red, NIR: Near Infra-Red, MIR: Medium IR, TIR: Thermal IR)</th><th colspan="4">Preliminary classification map output products: number of output spectral categories.</th></tr>
<tr><th>Fine semantic granularity</th><th>Intermediate semantic granularity</th><th>Coarse semantic granularity</th><th>Inter-sensor semantic granularity (*)</th></tr>
<tr><td>L-SIAM</td><td>7—B, G, R, NIR, MIR1, MIR2, TIR</td><td>96</td><td>48</td><td>18</td><td rowspan="6">33</td></tr>
<tr><td>S-SIAM</td><td>4—G, R, NIR, MIR1</td><td>68</td><td>40</td><td>15</td></tr>
<tr><td>AV-SIAM</td><td>4—R, NIR, MIR1, TIR</td><td>83</td><td>43</td><td>17</td></tr>
<tr><td>AA-SIAM</td><td>5—G, R, NIR, MIR1, TIR</td><td>83</td><td>43</td><td>17</td></tr>
<tr><td>Q-SIAM</td><td>4—B, G, R, NIR</td><td>61</td><td>28</td><td>12</td></tr>
<tr><td>D-SIAM</td><td>3—G, R, NIR</td><td>61</td><td>28</td><td>12</td></tr>
</table>

*Employed in sensor-independent bitemporal LCC detection.
Summary of input bands and output spectral categories reported in Table II.

V. CONCLUSION

In compliance with the QA4EO guidelines, the goal of this paper is to provide a theoretical comparison and an experimental quality assessment of two operational (ready-for-use) expert systems (prior knowledge-based nonadaptive decision trees) for automatic near real-time preattentional classification and segmentation of spaceborne/airborne MS images: the SIAM software product and the SPECL secondary product of the ATCOR commercial software toolbox. Rather than as standalone systems, these two alternative prior knowledge-based preclassifiers in operating mode are eligible for use in the preattentive vision first stage of a novel hybrid (combined deductive and inductive) RS-IUS architecture, proposed to the RS community in recent years [10]–[20].

For the sake of simplicity, this paper is split into two: Part 1—Theory, proposed herein, and Part 2—Experimental results, already published elsewhere [20].

The original contribution of the present Part 1 is threefold. First, it provides Part 2 with an interdisciplinary terminology and a theoretical background encompassing multiple disciplines, like philosophical hermeneutics, machine learning, artificial intelligence, computer vision, human vision and RS. Second, it highlights the relevant degrees of novelty of the ATCOR-SPECL and SIAM prior knowledge-based preclassifiers at the four levels of understanding of an information processing system, namely, system design, knowledge/information representation, algorithms and implementation. Third, it requires that a minimum set of community-agreed complementary independent metrological/statistically-based QIOs must be estimated from a RS-IUS in operating mode, to comply with the principles of statistics, the QA4EO guidelines [2] and the Committee on EO Satellites (CEOS) land product accuracy validation criteria [3]. In particular, sample QIs of the ATCOR-SPECL and SIAM prior knowledge-based preclassifiers, to be collected in Part 2 of this paper, must be: 1) statistically significant, i.e., provided with a degree of uncertainty in measurement, and 2) statistically valid (consistent), i.e., representative of the entire population being sampled, which requires the implementation of a probability sampling protocol [82], [83]. Noteworthy, these basic sample statistic requirements should not be considered either trivial or obvious. For example, they are almost never satisfied in the RS common practice. As a consequence, to date, QIOs of existing RS-IUSs, including mapping accuracy, in addition to degree of automation, efficiency, robustness, scalability, timeliness and costs, remain largely unknown in statistical terms.

The conclusion of the present Part 1 of this paper is that the proposed comparison of the ATCOR-SPECL and SIAM software products in operating mode, accomplished in Part 2, can be considered appropriate, well-timed and of potential interest to a wide portion of the RS community.

ACKNOWLEDGMENT

A. Baraldi would like to thank R. Capurro for his hospitality, patience, politeness, and open-mindedness. The authors would like to thank L. Boschetti for his support, who was at the University of Maryland, College Park, MD, USA, and is now with the College of Natural Resources, University of Idaho, Moscow, ID, USA, and the Editor-in-Chief, Associate Editor, and reviewers for their competence, patience, and willingness to help.

REFERENCES

[1] Group on Earth Observations (GEO). (2005, Feb. 16). *The Global Earth Observation System of Systems (GEOSS) 10-Year Implementation Plan* [Online]. Available: http://www.earthobservations.org/docs/10-Year%20Implementation%20Plan.pdf, accessed on Nov. 15, 2012.

[2] Group on Earth Observations (GEO)–Committee on Earth Observation Satellites (CEOS). (2010). *A Quality Assurance Framework for Earth Observation, Version 4.0* [Online]. Available: http://qa4eo.org/docs/QA4EO_Principles_v4.0.pdf, accessed on Nov. 15, 2012.

[3] Group on Earth Observations (GEO)-Committee on Earth Observation Satellites (CEOS) - CEOS Working Group on Calibration and Validation - Land Product Validation Subgroup. (2014, Aug. 19). *Land Product Validation Mission* [Online]. Available: http://lpvs.gsfc.nasa.gov/, accessed on Sep. 3, 2014.

[4] G. Gutman *et al.*, Eds., *Land Change Science*. Norwell, MA, USA: Kluwer, 2004.

[5] D. Marr, *Vision*. San Francisco, CA, USA: W.H. Freeman, 1982.

[6] W. Dorigo, R. Richter, F. Baret, R. Bamler, and W. Wagner, "Enhanced automated canopy characterization from hyperspectral data by a novel two step radiative transfer model inversion approach," *Remote Sens.*, vol. 1, pp. 1139–1170, 2009, doi: 10.3390/rs1041139.

[7] R. Richter and D. Schläpfer, *Atmospheric/topographic correction for satellite imagery, ATCOR-2/3 User Guide, Version 8.2.1*, DLR-IB 565-01/13. Wessling, Germany: DLR-German Aerospace Center, Feb. 2013 [Online]. Available: http://www.rese.ch/pdf/atcor3_manual.pdf, accessed on Sep. 3, 2014.

[8] R. Richter and D. Schläpfer, *Atmospheric/topographic correction for airborne imagery, ATCOR-4 User Guide, Version 6.2.1*, DLR-IB 565-02/13. Wessling, Germany, DLR-German Aerospace Center, Feb. 2013 [Online]. Available: http://www.rese.ch/pdf/atcor4_manual.pdf, accessed on Sep. 3, 2014.

[9] R. Richter, D. Schläpfer, and A. Hueni, "Recent developments in operational atmospheric and radiometric correction of hyperspectral imagery," in *Proc. 6th EARSeL SIG IS Workshop*, Tel Aviv, Israel, Mar. 16–19, 2009, pp. 1–6 [Online]. Available: http://www.earsel6th.tau.ac.il/∼earsel6/CD/PDF/earsel-PROCEEDINGS/3054%20Schl%20pfer.pdf, accessed on Sep. 3, 2014.

[10] 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 Observ. Remote Sens.*, vol. 2, no. 2, pp. 104–134, Jun. 2009.

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

[12] A. Baraldi *et al.*, "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.

[13] A. Baraldi *et al.*, "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.

[14] A. Baraldi *et al.*, "Corrections to '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," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 3, p. 1635, Mar. 2010.

[15] 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.

[16] A. Baraldi, T. Wassenaar, and S. Kay, "Operational performance of an automatic preliminary spectral rule-based decision-tree classifier of spaceborne very high resolution optical images," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 9, pp. 3482–3502, Sep. 2010.

[17] A. Baraldi, "Fuzzification of a crisp near-real-time operational automatic spectral-rule-based decision-tree preliminary classifier of multisource multispectral remotely sensed images," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 6, pp. 2113–2134, Jun. 2011.

[18] A. Baraldi and L. Boschetti, "Operational automatic remote sensing image understanding systems: Beyond geographic object-based and object-oriented image analysis (GEOBIA/GEOOIA)—Part 1: Introduction," *Remote Sens.*, vol. 4, no. 9, pp. 2694–2735, 2012, doi: 10.3390/rs4092694.

[19] A. Baraldi and L. Boschetti, "Operational automatic remote sensing image understanding systems: Beyond geographic object-based and object-oriented image analysis (GEOBIA/GEOOIA)—Part 2: Novel system architecture, information/knowledge representation, algorithm design and implementation," *Remote Sens.*, vol. 4, no. 9, pp. 2768–2817, 2012, doi: 10.3390/rs4092768.

[20] A. Baraldi, M. Humber, and L. Boschetti, "Quality assessment of preclassification maps automatically generated from spaceborne/airborne multi-spectral images by the Satellite Image Automatic Mapper and Atmospheric/Topographic Correction-Spectral Classification software products: Part 2—Experimental results," *Remote Sens.*, vol. 5, pp. 5209–5264, Oct. 2013.

[21] R. Capurro and B. Hjørland, "The concept of information," *Ann. Rev. Info. Sci. Tech.*, vol. 37, no. 1, pp. 343–411, 2003 [Online]. Available: http://www.capurro.de/infoconcept.html, accessed on 3 Sep. 2014.

[22] R. Capurro, "Hermeneutics and the phenomenon of information," in *Metaphysics, Epistemology, and Technology. Research in Philosophy and Technology*. Amsterdam, The Netherlands: JAI/Elsevier, 2000, vol. 19, pp. 79–85.

[23] E. Diamant, "Machine learning: When and where the horses went astray?" in *Machine Learning*, Y. Zhang, Ed. Rijeka, Croatia: InTech Open Access, 2010, pp. 1–18.

[24] V. Cherkassky and F. Mulier, *Learning From Data: Concepts, Theory, and Methods*. Hoboken, NJ, USA: Wiley, 1998.

[25] C. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pp. 379–423, 623–656, 1948.

[26] J. B. Adams *et al.*, "Classification of multispectral images based on fractions of endmembers: Application to land-cover change in the Brazilian Amazon," *Remote Sens. Environ.*, vol. 52, pp. 137–154, 1995.

[27] Yale Center for Earth Observation. (2009, Oct. 11). *South West Asia Project - Landcover Classification*. Yale University, Yale, NH, USA [Online]. Available: http://www.yale.edu/ceo/Projects/swap/landcover/Unsupervised_classification.htm, accessed on Sep. 3, 2014.

[28] J. Cihlar, R. Latifovic, and J. Beaubien, "A comparison of clustering strategies for unsupervised classification," *Can. J. Remote Sens.*, vol. 26, pp. 446–454, 2000.

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

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

[31] P. Mather, *Computer Processing of Remotely-Sensed Images—An Introduction*. Hoboken, NJ, USA: Wiley, 1994.

[32] 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.

[33] M. Baatz and A. Schäpe, "Multiresolution segmentation: An optimization approach for high quality multi-scale image segmentation," *J. Photogramm. Remote Sens.*, vol. 58, no. 3, pp. 12–23, 2000.

[34] M. Baatz, C. Hoffmann, and G. Willhauck, "Progressing from object-based to object-oriented image analysis," in *Object-Based Image Analysis—Spatial Concepts for Knowledge-Driven Remote Sensing Applications*, T. Blaschke, S. Lang and G. J. Hay, Eds. New York, NY, USA: Springer, 2008, ch. 1.4, pp. 29–42.

[35] S. Tatiraju and A. Mehta, "Image Segmentation using k-means clustering, EM and normalized cuts," Univ. California, Irvine, CA, USA, Tech. Rep., 2008, pp. 1–7 [Online]. Available: http://www.ics.uci.edu/∼dramanan/teaching/winter08/projects/avim˙report.pdf, accessed on Sep. 3, 2014.

[36] N. R. Pal and S. K. Pal, "A review on image segmentation techniques," *Pattern Recognit.*, vol. 26, no. 9, pp. 1277–1294, 1993.

[37] A. Baraldi and E. Alpaydin, "Constructive feedforward ART clustering networks—Part I," *IEEE Trans. Neural Netw.*, vol. 13, no. 3, pp. 645–661, May 2002.

[38] A. Baraldi and E. Alpaydin, "Constructive feedforward ART clustering networks—Part II," *IEEE Trans. Neural Netw.*, vol. 13, no. 3, pp. 662–677, May 2002.

[39] B. Fritzke, "Some competitive learning methods," Bochum, Germany: Ruhr-Universität Bochum, 1997 [Online]. Available: http://www.ki.inf.tu-dresden.de/∼fritzke/JavaPaper

[40] E. Backer and A. K. Jain, "A clustering performance measure based on fuzzy set decomposition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-3, no. 1, pp. 66–75, Jan. 1981.

[41] C. Persello and L. Bruzzone, "A novel protocol for accuracy assessment in classification of very high resolution images," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 3, pp. 1232–1244, Mar. 2010.

[42] J. Hadamard, "Sur les problemes aux derivees partielles et leur signification physique," *Princeton Univ. Bull.*, vol. 13, pp. 49–52, 1902.

[43] N. Hunt and S. Tyrrell. (2004, Jun. 10). *Stratified Sampling*. Coventry University, U.K. [Online]. Available: http://nestor.coventry.ac.uk/∼nhunt/meths/strati.html, accessed on 3 Sep. 2014.

[44] P. Zamperoni, "Plus ça va, moins ça va," *Pattern Recognit. Lett.*, vol. 17, pp. 671–677, 1996.

[45] S. Liang, *Quantitative Remote Sensing of Land Surfaces*. Hoboken, NJ, USA: Wiley, 2004.

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

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

[48] T. F. Cootes and C. J. Taylor, *Statistical Models of Appearance for Computer Vision, Imaging Science and Biomedical Engineering*. Manchester, U.K.: Univ. Manchester, 2004.

[49] D. K. Hall, G. A. Riggs, V. V. Salomonson, N. E. Di Girolamo, and K. J. Bayr, "MODIS snow-cover products," *Remote Sens. Environ.*, vol. 83, pp. 181–194, 2002.

[50] Y. Luo, A. P. Trishchenko, and K. V. Khlopenkov, "Developing clear-sky, cloud and cloud shadow mask for producing clear-sky composites at 250-meter spatial resolution for the seven MODIS land bands over Canada and North America," *Remote Sens. Environ.*, vol. 112, pp. 4167–4185, 2008.

[51] A. K. Jain, "Knowledge-based segmentation of Landsat images," *IEEE Trans. Geosci. Remote Sens.*, vol. 29, no. 2, pp. 222–232, Mar. 1991.

[52] M. Avcia and Z. Akyurek, "A hierarchical classification of Landsat TM imagery for landcover mapping," in *Proc. Int. Soc. Photogramm. Remote Sens. ISPRS XXXV Congr. WG IV/6*, Istanbul, Turkey, Jul. 12–23, 2004, pp. 1–6 [Online]. Available: http://www.isprs.org/proceedings/XXXV/congress/comm4/papers/402.pdf, accessed on Sep. 3, 2014.

[53] L. Moller-Jensen, "Classification of urban land cover based on expert systems, object models and texture," *Environ. Urban Syst.*, vol. 21, no. 3/4, pp. 291–302, 1997.

[54] A. K. Skidmore, "An expert system classifies Eucalypt forest types using TM data and a digital terrain model," *Photogramm. Eng. Remote Sens.*, vol. 55, no. 10, pp 1449–1464, 1989.

[55] S. Lang, "Object-based image analysis for remote sensing applications: Modeling reality—Dealing with complexity," in *Object-Based Image Analysis–Spatial Concepts for Knowledge-Driven Remote Sensing Applications*, T. Blaschke, S. Lang and G. J. Hay, Eds. New York, NY, USA: Springer, 2008, ch. 1.1, pp. 3–27.

[56] G. J. Hay and G. Castilla, "Object-based image analysis: Strengths, weaknesses, opportunities and threats (SWOT)," in *Proc. 1st Int. Conf. Object-Based Image Anal. (OBIA)*, Salzburg, Austria, Jul. 2006, pp. 1–3.

[57] G. J. Hay and G. Castilla, "Geographic object-based image analysis (GEOBIA): A new name for a new discipline," in *Object-Based Image Analysis: Spatial Concepts for Knowledge-Driven Remote Sensing Applications*, T. Blaschke, S. Lang and G. J. Hay, Eds. New York, NY, USA: Springer, 2008, ch. 1.4, pp. 81–92.

[58] Definiens Imaging GmbH, *eCognition Elements User Guide 4*. Munich, Germany: Definiens Imaging GmbH, 2004.

[59] Definiens AG, *Developer 8 Reference Book*. Munich, Germany: Definiens AG, 2011.

[60] Trimble. (2014). *eCognition Developer 9.0* [Online]. Available: http://www.ecognition.com/products/ecognition-developer, accessed on Sep. 3, 2014.

[61] 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. 2013.

[62] 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.

[63] R. R. Irish, J. L. Barker, S. N. Goward, and T. Arvidson, "Characterization of the Landsat-7 ETM+ automated cloud-cover assessment (ACCA) algorithm," *Photogramm. Eng. Remote Sens.*, vol. 72, no. 10, pp. 1179–1188, Oct. 2006.

[64] Q. Iqbal and J. K. Aggarwal, "Image retrieval via isotropic and anisotropic mappings," in *Proc. IAPR Workshop Pattern Recognit. Inf. Syst.*, Setubal, Portugal, Jul. 2001, pp. 34–49.

[65] M. F. Goodchild, M. Yuan, and T. J. Cova, "Towards a general theory of geographic representation in GIS," *Int. J. Geogr. Inf. Sci.*, vol. 21, no. 3, pp. 239–260, 2007.

[66] P. Lüscher, D. Burghardt, and R. Weibel, "Ontology-driven enrichment of spatial databases," in *Proc. 10th ICA Workshop Gen. Multiple Represent.*, Moscow, Russia, Aug. 2007, pp. 1–13 [Online]. Available: http://www.geo.uzh.ch/∼luescher/publications/luescher_genws2007.pdf, accessed on Sep. 3, 2014.

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

[68] P. Gouras, "Color vision," in *Principles of Neural Science*, E. Kandel and J. Schwartz, Eds. Norwalk, CT, USA: Appleton and Lange, 1991, pp. 467–479.

[69] E. R. Kandel, "Perception of motion, depth and form," in *Principles of Neural Science*, E. Kandel and J. Schwartz, Eds. Norwalk, CT, USA: Appleton and Lange, 1991, pp. 441–466.

[70] M. Carandini and D. Ferster, "Membrane potential and firing rate in cat primary visual cortex," *J. Neuroscience*, vol. 20, pp. 470–484, 2000.

[71] J. Mutch and D. Lowe, "Object class recognition and localization using sparse features with limited receptive fields," *Int. J. Comput. Vis.*, vol. 80, pp. 45–57, 2008.

[72] T. N. Wiesel and D. H. Hubel, "Spatial and chromatic interactions in the lateral geniculate body of the rhesus monkey," *J. Neurophys.*, vol. 29, pp. 1115–1156, 1966.

[73] A. Jain and G. Healey, "A multiscale representation including opponent color features for texture recognition," *IEEE Trans. Image Proc.*, vol. 7, no. 1, pp. 124–128, Jan. 1998.

[74] T. Martinetz and K. Schulten, "Topology representing networks," *Neural Networks*, vol. 7, no. 3, pp. 507–522, 1994.

[75] S. P. Vecera and M. J. Farah, "Is visual image segmentation a bottom-up or an interactive process?" *Percept. Psychophys.*, vol. 59, 1280–1296, 1997.

[76] J. D. McCafferty, *Human and Machine Vision, Computing Perceptual Organization*. Chichester, U.K.: Ellis Horwood Limited, 1990.

[77] J. Yang and R. S. Wang, "Classified road detection from satellite images based on perceptual organization," *Int. J. Remote Sens.*, vol. 28, pp. 4651–4669, 2007.

[78] Wikipedia. (2014, Sep. 1). *Big data* [Online]. Available: http://en.wikipedia.org/wiki/Big_data, accessed on Sep. 3, 2014.

[79] L. Alparone, S. Baronti, A. Garzelli, and F. Nencini, "A global quality measurement of pan-sharpened multispectral imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 1, no. 4, pp. 313–317, Oct. 2004.

[80] A. Baraldi, L. Boschetti, and M. Humber, "Probability sampling protocol for thematic and spatial quality assessments of classification maps generated from spaceborne/airborne very high resolution images," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 1, Part: 2, pp. 701–760, Jan. 2014.

[81] Y. Y. Boykov and M. P. Jolly, "Interactive graph cuts for optimal boundary and region segmentation of objects in N-D images," in *Proc. Int. Conf. Comput. Vis.*, Vancouver, BC, Canada, Jul. 2001, vol. I, pp. 105–112.

[82] S. V. Stehman and R. L. Czaplewski, "Design and analysis for thematic map accuracy assessment: Fundamental principles," *Remote Sens. Environ.*, vol. 64, pp. 331–344, 1998.

[83] W. S. Overton and S. V. Stehman, "The Horvitz–Thompson theorem as a unifying perspective for probability sampling: With examples from natural resource sampling," *Amer. Statistician*, vol. 49, no. 3, pp. 261–268, Aug. 1995.

[84] M. A. Wulder, J. G. Masek, W. B. Cohen, T. R. Loveland, and C. E. Woodcock, "Opening the archive: How free data has enabled the science and monitoring promise of Landsat," *Remote Sens. Environ.*, vol. 122, pp. 2–10, July 2012.

[85] ENVI 4.3, User Manual, ITT Industries, Inc., 2006.

[86] E. Vermote, D. Tanré, J. L. Deuzé, M. Herman, and J. J. Morcrette. (1997, Jul.). *Second Simulation of the Satellite Signal in the Solar Spectrum (6S), 6S User Guide Version 2* [Online]. Available: http://www.rsgis.ait.ac.th/∼honda/textbooks/advrs/6smanv2.0_P1.pdf, accessed on Jan. 10, 2012.

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

[88] 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.

[89] P. S. Chavez, "An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data," *Remote Sens. Environ.*, vol. 24, pp. 459–479, 1988.

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

[91] R. Richter, DLR-German Aerospace Center, personal communication to A. Baraldi.

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

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

[94] Z. Wang and A. C. Bovik, "A universal image quality index," *IEEE Signal Process. Lett.*, vol. 9, no. 3, pp. 81–84, Mar. 2002.

[95] G. Schaepman-Strub, M. E. Schaepman, T. H. Painter, S. Dangel, and J. V. Martonchik, "Reflectance quantities in optical remote sensing-definitions and case studies," *Remote Sens. Environ.*, vol. 103, pp. 27–42, 2006.

[96] European Commission Joint Research Center (JRC)–Institute for Environment and Sustainability (IES). *IMAGE2000 and CORINE Land Cover (CLC) 2000*. Ispra, Italy: JRC–IES, 2005, EUR 21757, ISBN 92-894-9862-5.

[97] European Commission Joint Research Center, *The IMAGE-2006 Mosaic Project*, EC-JRC: Ispra, Italy, EUR 23755, 2008 [Online]. Available: http://publications.jrc.ec.europa.eu/repository/bitstream/111111111/22218/1/lbna23755enn.pdf, accessed on Sep. 3, 2014.

[98] 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 N. Bahram, Eds. Boca Raton, FL, USA: CRC Press, 1992, pp. 309–327.

[99] H. R. Wilson and J. R. Bergen, "A four mechanism model for threshold spatial vision," *Vis. Res.*, vol. 19, no. 1, pp. 19–32, 1979.

[100] H. du Buf and J. Rodrigues, "Image morphology: From perception to rendering," in *IMAGE—Computational Visualistics and Picture Morphology*, 2007 [Online]. Available: http://w3.ualg.pt/∼jrodrig/papers_pdf/IMAGE07.pdf, accessed on Aug. 16, 2014.

[101] D. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, 2004.

[102] J. van de Weijer, C. Schmid, J. Verbeek, and D. Larlus, "Learning color names for real-world applications," *IEEE Trans. Image Process.*, vol. 18, no. 7, pp. 1512–1523, Jul. 2009.

[103] R. Benavente, M. Vanrell, and P. Baldrich, "Parametric fuzzy sets for automatic color naming," *J. Opt. Soc. Amer. A*, vol. 25, pp. 2582–2593, 2008.

[104] A. Plaza, J. Plaza, and G. Martin, "Spatial-spectral endmember extraction from hyperspectral imagery using multi-band morphology and volume optimization," in *Proc. 16th IEEE Int. Conf. Image Process. (ICIP)*, Cairo, Egypt, Nov. 2009, pp. 3721–3724.

Andrea Baraldi received the Laurea (M.S.) degree in electronic engineering from the University of Bologna, Bologna, Italy, in 1989, and a Master degree in software engineering from the University of Padova, Padova, Italy, and Purdue University, West Lafayette, IN, USA, in 1994.

From 1989 to 1990, he worked as a Research Associate at the Centro di Studio per l'Interazione Operatore–Calcolatore, National Research Council (CNR), Bologna, Italy, and served in the army with the Istituto Geografico Militare, Florence, Italy. He worked as a Consultant with the European Space Agency–European Space Research Institute, Frascati, Italy, from 1991 to 1993. From 1997 to 1999 he was a PostDoctoral Fellow in Artificial Intelligence with the International Computer Science Institute, Berkeley, CA, USA. He worked as a PostDoctoral Researcher with the Global Vegetation Monitoring Unit, the Institute for Environment and Sustainability (IES), the European Commission Joint Research Centre (JRC), Ispra, Italy, from 2000 to 2002. From 2005 to 2009, he was with the IES-Spatial Data Infrastructure, JRC. In 2009, he founded Baraldi Consultancy in Remote Sensing, a one-man company located in Modena, Italy. Since his master thesis until 2010 he continued his collaboration with the Istituto di Scienze dell'Atmosfera e del Clima, CNR, Bologna, and the Istituto di Studi di Sistemi Intelligenti per l'Automazione, CNR, Bari, Italy. From 2010 to 2013, he joined the Department of Geographical Studies, University of Maryland, College Park, MD, USA, where, to date, he is an Adjunct Associate Professor. In 2014, he was appointed with a Senior Scientist Fellowship at the German Aerospace Center (DLR), Oberpfaffenhofen, Germany. From June 2014 to October 2104, he was a Visiting Scientist at the Department of Geoinformatics—Z_GIS, University of Salzburg, Salzburg, Austria. His research interests include image preprocessing and understanding, with special emphasis on the development of operational automatic hierarchical multisource multiresolution spaceborne/airborne image understanding systems, consistent with human vision.

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

Michael L. Humber received the M.P.S. degree in geospatial information sciences in 2013, and is pursuing the Ph.D. degree at the Department of Geographical Sciences, University of Maryland, College Park, MD, USA.

He is a Research Assistant with the University of Maryland. Since 2010, he has been employed with the Department of Geographical Sciences Research Laboratory. His research interests include processing and validation of the MODIS Burned Area Products, validation of maps generated from moderate to very high resolution imagery, global agricultural monitoring, and web mapping in the context of facilitated volunteered geographic information (VGI) applications.

Quality Assessment of Preclassification Maps Generated From Spaceborne/Airborne Multispectral Images by the *Satellite Image Automatic Mapper* and *Atmospheric/Topographic Correction-Spectral Classification* Software Products: Part 1—Theory

Andrea Baraldi and Michael L. Humber

***Abstract*—In compliance with the Quality Assurance Framework for Earth Observation (QA4EO) guidelines, the goal of this paper is to provide a theoretical comparison and an experimental quality assessment of two operational (ready-for-use) expert systems (prior knowledge-based nonadaptive decision trees) for automatic near real-time preattentional classification and segmentation of spaceborne/airborne multispectral (MS) images: the Satellite Image Automatic Mapper™ (SIAM™) software product and the Spectral Classification of surface reflectance signatures (SPECL) secondary product of the Atmospheric/Topographic Correction™ (ATCOR™) commercial software toolbox. For the sake of simplicity, this paper is split into two: Part 1—Theory, presented herein, and Part 2—Experimental results, already published elsewhere. The main theoretical contribution of the present Part 1 is threefold. First, it provides the published Part 2 with an interdisciplinary terminology and a theoretical background encompassing multiple disciplines, such as philosophical hermeneutics, machine learning, artificial intelligence, computer vision, human vision, and remote sensing (RS). Second, it highlights the several degrees of novelty of the ATCOR-SPECL and SIAM deductive preliminary classifiers (preclassifiers) at the four levels of abstraction of an information processing system, namely, system design, knowledge/information representation, algorithms, and implementation. Third, the present Part 1 requires the experimental Part 2 to collect a minimum set of complementary statistically independent metrological quality indicators (QIs) of operativeness (QIOs), in compliance with the QA4EO guidelines and the principles of statistics. In particular, sample QIs are required to be: 1) statistically significant, i.e., provided with a degree of uncertainty in measurement; and 2) statistically valid (consistent), i.e., representative of the entire population being sampled, which requires the implementation of a probability sampling protocol. Largely overlooked by the RS community, these sample QI requirements are almost never satisfied in the RS common practice. As a consequence, to date, QIOs of existing RS image understanding systems (RS-IUSs), including thematic map accuracy, remain largely unknown in statistical terms. The conclusion of the present Part 1 is that the proposed comparison of the two alternative ATCOR-SPECL and SIAM prior knowledge-based preclassifiers in operating mode, accomplished in the Part 2, can be considered appropriate, well-timed, and of potential interest to a large portion of the RS readership.**

***Index Terms*—Attentive vision, degree of uncertainty in measurement, land cover classification taxonomy, preattentive vision, preliminary classification, probability sampling, quality indicator (QI), radiometric calibration, spectral category, spectral mixture analysis.**

Manuscript received January 05, 2014; accepted August 06, 2014. This work was supported by the National Aeronautics and Space Administration under Grant NNX07AV19G issued through the Earth Science Division of the Science Mission Directorate.

A. Baraldi was with the University of Maryland, College Park, MD 20742 USA. He is now with Baraldi Consultancy in Remote Sensing, Modena 41121, Italy (e-mail: andrea6311@gmail.com).

M. L. Humber is with the Department of Geographical Sciences, University of Maryland, College Park, MD 20742 USA (e-mail: mhumber@umd.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JSTARS.2014.2349932

I. Introduction

One visionary goal of the remote sensing (RS) community is to develop information processing systems capable of automatically transforming, without user interactions, large-scale multisource multiresolution Earth observation (EO) image databases into "operational, comprehensive, and timely knowledge/information products" [1]–[3], at spatial extents ranging from local to global [4]. The Quality Assurance Framework for EO (QA4EO) guidelines [2], [3], conceived by the international Group on EOs (GEO)-Committee on EO Satellites (CEOS), comprise an extensive formulation of this ambitious goal. For example, the ongoing GEO Global EO System of Systems (GEOSS) implementation plan for years 2005–2015 incorporates the QA4EO guidelines to build a global public infrastructure that allows "the provision of and access to the Right (geospatial) Information, in the Right Format, at the Right Time, to the Right People, to Make the Right Decisions" [1].

To pave the way for the design and implementation of a novel generation of automatic RS image understanding systems (RS-IUSs) in compliance with the QA4EO guidelines [2], [3], this paper provides a theoretical comparison and an experimental quality assessment of two operational (ready-for-use) expert systems (prior knowledge-based nonadaptive decision trees) for automatic near real-time preliminary classification (preclassification [5]) and segmentation of spaceborne/airborne EO multispectral (MS) images: the spectral classification of surface reflectance signatures (SPECL) software module and the Satellite Image Automatic Mapper (SIAM) software product. The former is implemented as a nonvalidated secondary product within the popular Atmospheric/Topographic Correction (ATCOR)-2/3/4 commercial software toolbox [6]–[9]. The latter has been presented in recent years in the RS literature [10]–[19], where enough information is provided for the SIAM implementation to be reproduced [11], [17].

Rather than being considered as standalone software products, the two alternative ATCOR-SPECL and SIAM expert systems for automatic near real-time preclassification and segmentation of multisource MS images are eligible for use in the preattentive vision first stage of a novel generation of automatic *hybrid* (combined deductive and inductive) RS-IUS implementations [10]–[20].

For the sake of simplicity, this paper is split into two: the Part 1—Theory, presented herein, and the Part 2—Experimental results, already published elsewhere [20]. The main theoretical contribution of the present Part 1 is threefold. First, it provides the Part 2 with an interdisciplinary terminology and a theoretical background encompassing multiple disciplines, such as philosophical hermeneutics, machine learning, artificial intelligence, computer vision, human vision, and RS. Hence, Part 1 is provided with a relevant survey value. Second, it highlights the relevant degrees of novelty of the ATCOR-SPECL and SIAM prior knowledge-based preclassifiers at the four levels of abstraction of an information processing system, namely, system design, knowledge/information representation, algorithms, and implementation. Third, the present Part 1 requires the experimental Part 2 to collect a minimum set of complementary independent metrological/statistically-based quality indicators (QIs) of operativeness (QIOs), in compliance with the QA4EO guidelines and the principles of statistics. In particular, sample QIs are required to be: 1) statistically significant, i.e., provided with a degree of uncertainty in measurement and 2) statistically valid (consistent), i.e., representative of the entire population being sampled, which requires the implementation of a probability sampling protocol. Largely overlooked by the RS community, these sample QI requirements are almost never satisfied in the RS common practice. As a consequence, to date, QIOs of existing RS-IUSs, including thematic map accuracy, remain largely unknown in statistical terms. The conclusion of the present Part 1 is that the proposed comparison of the two alternative ATCOR-SPECL and SIAM prior knowledge-based preclassifiers in operating mode, accomplished in the Part 2, can be considered appropriate, well-timed, and of potential interest to a large portion of the RS readership.

The rest of the present Part 1 is organized as follows. Section II presents an interdisciplinary terminology and a theoretical background useful for the understanding of the experimental Part 2. Problem recognition and opportunity identification are discussed in Section III. In Section IV, the two alternative ATCOR-SPECL and SIAM preclassification expert systems are compared at the four levels of abstraction of an information processing system. Conclusion of this theoretical contribution is reported in Section V.

II. Interdisciplinary Terminology and Problem Background

According to Section I, the goal of the experimental Part 2 of this paper, published elsewhere [20], is to pursue a statistically significant and statistically consistent quality assessment of the ATCOR-SPECL and SIAM deductive preclassification software products in operating mode, eligible for use in the preattentive vision first stage of a hybrid RS-IUS architecture [20]. Introduced by Section I, terms such as "statistically significant" QI, "statistically consistent" probability sampling, "QIOs of an information processing system in operating mode," "quality assessment of a preclassification map," "deductive preclassification," "preattentive/attentive vision," "deductive/inductive/hybrid inference," and "data/information/knowledge" are defined explicitly and unambiguously in this section, based on a multidisciplinary approach. To be employed in the rest of the present Part 1 and in the Part 2, the proposed interdisciplinary terminology provides this paper with a significant survey value.

A. Quantitative and Qualitative Concepts of Information

Philosophical hermeneutics refers to the theory of knowledge and the practice, art or science of (text) interpretation and explanation. According to philosophical hermeneutics [21], [22], the impact upon computer science, information technology (IT), artificial intelligence and machine learning of existing different quantitative and qualitative concepts of information, embedded in more or less explicit information theories, appears largely underestimated. This means that fundamental questions—like: When do (subsymbolic) data become (symbolic) information [23]? When does vision go symbolic [5]? Should traditional information retrieval be called document retrieval [21], [22]?—appear largely overlooked and, as a consequence, far from being answered.

In accordance with philosophical hermeneutics, the fundamental concepts of *numerical data*, *quantitative information*, *qualitative information* and *knowledge* are defined hereafter [21], [22].

1) Numerical data, sensory data, quantitative data, observational data are considered synonyms of "true facts" [24]. *Sensory data are provided, per se, with no semantics at all* [23], i.e., observational data are always subsymbolic (unlabeled).
2) Subsymbolic, quantitative, unequivocal "*information-as-thing*" is, according to the Shannon theory of communication [25], an object or a thing (e.g., number of bits and number of words in a document) irrespective of its meaning. This makes the information exchange between a sender and a receiver unequivocal (context independent) and, therefore, easier to deal with than when meaning is involved in the communication process [18], [19], [21], [22].
3) Symbolic, qualitative, equivocal "*information-as-(an intepretation)process*," i.e., information as interpreted data, is, in the words of philosophical hermeneutics, symbolic information always related to "a receiver's beliefs,

desires and background knowledge" [21], [22]: the meaning of a message is always context-dependent, depending on (changing with) the inquirer (user, knower, receiver, cognitive agent) in charge of the message interpretation. For example, Adams *et al.* underline that land cover (LC) "class names are selected to have significance to an observer in the field and in the context of a given study" [26].

4) "Knowledge" is strictly related to the concept of "*information-as-(an intepretation)process*," such that "there is no knowledge without both an object of knowledge and a knowing subject." [21], [22]. Hence, "information-as-(an interpretation)process" and "knowledge" can be considered as synonyms. A well-known example of equivocal (subjective, context-dependent) interpretation process is the so-called "fusion of ontologies" or "fusion of thematic map legends" [21], [22], occurring when two thematic maps of the same geographic area, but featuring different map legends, must be compared. In other words, it is reasonable to expect that two independent domain experts required to harmonize (reconcile) two thematic map legends may fulfill their (inherently equivocal) interpretation processes with different inter-vocabulary mapping functions.

Noteworthy, *the complementary concepts of information-as-(an interpretation)process and information-as-thing apply one-to-one to the dual concepts of (equivocal, qualitative, symbolic) categorical (nominal) variables and (unequivocal, quantitative, subsymbolic) continuous/discrete scalar/vector variables* (e.g., biophysical variables, such as leaf area index and biomass), *to be estimated from sensory data* [18], [19], [47]. To conclude, the following terms can be considered as nontrivial synonyms.

1) Symbolic, semantic, cognitive, categorical, ordinal, nominal, qualitative, subjective, equivocal. For example, (discrete and symbolic) categorical variable.
2) Subsymbolic, sensory, numerical, nonsemantic, quantitative, objective, unequivocal. For example, (subsymbolic) continuous or discrete sensory variable.

For example, according to the terminology proposed herein, the two ATCOR-SPECL and SIAM prior knowledge-based preclassifiers, to be assessed and compared in the Part 2 [20], automatically transform (subsymbolic quantitative) MS images (2-D data) into a (symbolic qualitative) categorical variable, whose values belong to a discrete and finite legend of (semantic) concepts.

B. Inductive, Deductive, and Hybrid Inference Systems, Either Subsymbolic or Symbolic, Investigated by the Machine Learning, Artificial Intelligence, and RS Disciplines

This section introduces expressions like inductive, deductive and hybrid inference system, either subsymbolic or symbolic (refer to Section II-A), depending on whether the inference system deals with, respectively, subsymbolic variables, either continuous or discrete, or (symbolic and discrete) categorical (nominal) variables. The specialization capability of this terminology is far superior to that of expressions traditionally used or misused by the RS community, such as supervised or unsupervised data learning. For example, an expression such as "unsupervised classification" is widely adopted by the RS community to mean either "unsupervised data clustering" or "automatic classification," e.g., see [27] and [28]. Unfortunately, according to the machine learning literature, this expression is a typical contradiction of terms because: 1) "unsupervised," e.g., unsupervised data, refers to "unlabeled," e.g., unlabeled data, rather than "without user's supervision," i.e., "unsupervised" does not mean "automatic" and 2) sensory data are provided with no semantics at all (refer to Section II-A), i.e., observational data are always, *per se*, unsupervised (unlabeled), while, by definition, classified data are always supervised (labeled) data, where data labels belong to a discrete and finite taxonomy of (semantic) concepts [23], [24], [29].

Hereafter, the concepts of inductive, deductive and hybrid inference system, either subsymbolic or symbolic, are discussed in detail.

There are two classical types of inference (learning), known as: 1) *induction*, progressing from particular cases (e.g., true facts and training data samples) to a general estimated dependency or model, and 2) *deduction*, progressing from a general model (e.g., a physical model-based equation) to particular cases (e.g., output values) [24]. Inductive inference is the basis of the machine learning discipline [24], [29]. Deductive inference is the main focus of interest of traditional artificial intelligence [24], [29]–[31].

The following terms are nontrivial synonyms of deductive inference and become interchangeable in the rest of this work [18], [19]: (subsymbolic or symbolic) deductive inference, deductive learning, top-down inference system, coarse-to-fine inference, driven-by-knowledge inference, learning-by-rules, physical model, prior knowledge-based decision system, rule-based system, expert system, syntactic inference, and syntactic pattern recognition.

The following terms are nontrivial synonyms of inductive inference [18], [19]: (subsymbolic or symbolic) inductive inference, inductive learning from either labeled (supervised) or unlabeled (unsupervised) data, bottom-up inference, fine-to-coarse inference, driven-without-knowledge (knowledge-free) inference, learning-from-examples, statistical model.

For the sake of completeness, some well-known examples of inductive and deductive inference systems, presented in the computer vision, machine learning and/or RS literature, are listed as follows.

1) In the computer vision literature, image segmentation algorithms are typical examples of subsymbolic inductive inference systems for unlabeled data learning [32]–[36].
2) In the machine learning literature, unsupervised (unlabeled) data learning algorithms are either vector data quantizers (e.g., the well-known k-means data quantization algorithm, improperly called k-means data clustering algorithm), probability density function estimators or unlabeled data clustering algorithms [15], [24], [29], [37]–[40]. Inductive supervised (labeled) data learning systems are either: 1) symbolic (classifiers), e.g., artificial neural network classifiers, support vector machine

classifiers [41], nearest-neighbor classifiers, adaptive decision-tree classifiers, and radial basis function networks for classification [24], [29] or 2) subsymbolic, suitable for function regression, e.g., radial basis function networks for function regression [24], [29].

3) In the RS literature [24], [29], a typical example of subsymbolic inductive inference system is principal component analysis; a popular example of subsymbolic deductive inference system is tasseled cap transformation.

The machine learning literature clearly acknowledges that all inductive data learning problems are inherently ill-posed in the Hadamard sense [42]. According to Hadamard, mathematical or statistical models of physical phenomena are defined as well-posed (respectively, ill-posed) when they satisfy (respectively, do not satisfy at least one of) the following requirements [42]: 1) a solution exists, 2) the solution is unique, and 3) the solution's behavior hardly changes when there is a slight change in the initial condition. In the words of Mulier and Cherkassky: "*induction amounts to forming generalizations from particular true facts. This is an inherently difficult (ill-posed) problem and its solution requires a priori knowledge in addition to data*" [24] (p. 39). Hence, to become better posed (conditioned) for numerical treatment, any inductive data learning algorithm requires an *a priori* knowledge base (deductive inference approach) to avoid starting from scratch when looking at input sensory data [10]–[19]. This conclusion complies with the well-known statistical principle of *stratification*, equivalent to the divide-and-conquer (*dividi et impera*) problem solving approach [29], to be enforced upon statistical systems. The advantage of a stratified statistical system is that it "*will always achieve greater precision (than its nonstratified counterpart), 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*" [43].

On one hand, well-known limitations of statistical (bottom-up inference) systems in common practice are that they are inherently semiautomatic and site-specific [18], [45]. On the other hand, typical drawbacks of physical (top-down inference) models are that [18]: 1) in general, it takes a long time for human experts to learn physical laws of the real-world-through-time and tune physical models, 2) physical models suffer from an intrinsic lack of flexibility, i.e., decision rules do not adapt to changes in the input data format and users' needs, hence their knowledge base may soon become obsolete, and 3) physical models suffer from an intrinsic lack of scalability, in particular rule-based systems are impractical for complex problems [30].

There is an ongoing multidisciplinary debate about a claimed inadequacy of scientific disciplines such as computer vision, artificial intelligence, and machine learning, whose origins date back to the late 1950s, in the provision of operational solutions to their ambitious cognitive objectives [23], [44]. This claim may mean that, if they are not combined, inductive and deductive inference approaches show intrinsic weaknesses in operational use, irrespective of implementation [18]. As a consequence, to outperform existing deductive and inductive inference systems whose drawbacks are well known, a novel trend in recent literature aims at developing hybrid inference systems for retrieval of subsymbolic variables (e.g., leaf area index, LAI) or symbolic variables (e.g., LC and LC change (LCC) classes) from sensory data (e.g., optical imagery) [45]–[48]. By definition, *hybrid inference systems, either subsymbolic or symbolic, combine both statistical and physical models to take advantage of the unique features of each and overcome their shortcomings* [46], [47]. For example, in the foreword of the seminal book by Nagao and Matsuyama [47], published in 1980 (oldies, but goldies), it is written: "The work described here is a deep *unification and synthesis of the two fundamental approaches to pattern recognition: numerical (also known as 'statistical') and structural ('linguistic,' 'syntactic').*"

Noteworthy, physical model-based inference systems as well as hybrid models require as input observational data provided with a physical meaning, i.e., sensory data provided with a physical unit of measure, e.g., RS imagery radiometrically calibrated into top-of-atmosphere (TOA) radiance or TOA reflectance values [10]. On the other hand, statistical systems can be input with any sort of numerical data, irrespective of their physical meaning, if any. This is tantamount to saying that, whereas *dimensionless sensory data, provided with no physical unit of measure, are eligible for use as input to statistical models exclusively*, on the contrary, *numerical data provided with a physical unit of measure can be input to both physical and statistical models.*

For the sake of completeness, let us review some additional examples of inductive, deductive and hybrid RS-IUS instances proposed in recent years in the RS literature. A large family of one-stage one-pass (noniterative) prior knowledge-based (static, nonadaptive to input data) decision-tree (pre)classifiers (symbolic expert systems) has been proposed, starting from the 1970 s, as a legacy of traditional artificial intelligence [49], [50], [51]–[54]. For example, in [50] (p. 4176), a one-stage physical model-based RS-IUS, see Fig. 1(a), consists of a hierarchy of five pixel-specific prior knowledge-based spectral rules proposed to detect six land surface types, namely, "vegetated lands," "nonvegetated lands," "snow/ice," "water bodies," "clouds," and "cloud shadows," in radiometrically calibrated 500 m resolution moderate resolution imaging spectroradiometer (MODIS) images. In 30 m resolution Landsat images, a one-stage deductive RS-IUS, consisting of a hierarchy of per-pixel prior knowledge-based spectral rules, detects LC classes "water," "coniferous forest," "deciduous forest," "agricultural areas," "grassland," "urban areas," and "roads" [52]. In recent years, prior knowledge-based decision-tree classifiers are employed per image-object at an attentive vision second stage, in series with an inductive image segmentation first stage, like in the popular two-stage noniterative Geographic Object-Based Image Analysis (GEOBIA) system architecture, see Fig. 1(b), and in the three-stage iterative Geographic Object-Observation Image Analysis (GEOOIA) system design [32]–[34], [55]–[60]. The former is a special case of the latter, i.e., GEOBIA $\subseteq$ GEOOIA, where both GEOBIA and GEOOIA share a statistical model-based subsymbolic image segmentation first stage. Alternative to GEOBIA/GEOOIA systems, an original two-stage hybrid RS-IUS architecture is proposed by Shackelford and Davis [61], [62]. It comprises an image-object-based expert system for second-stage decision-tree classification in series with a first-stage pixel-based

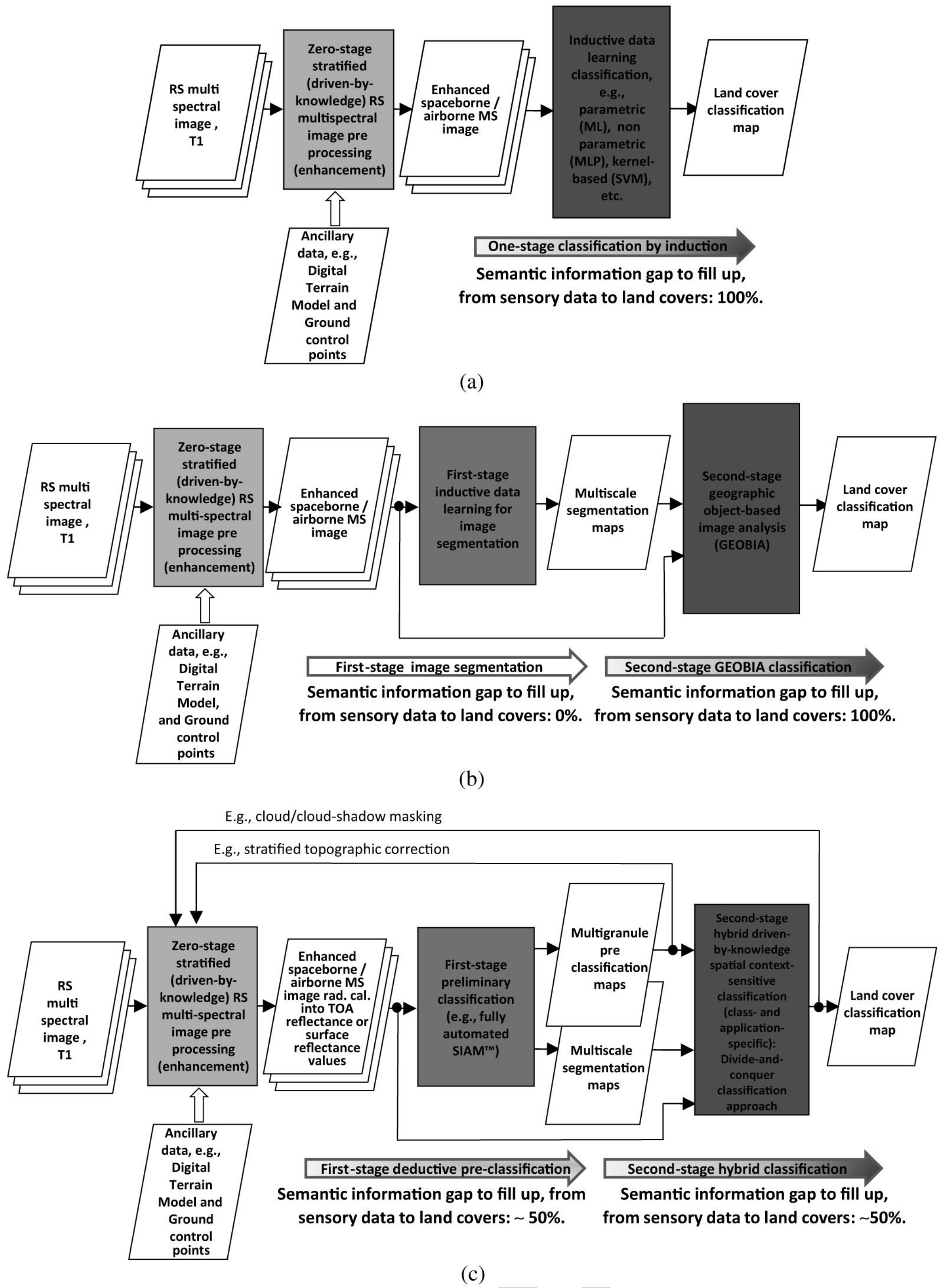

Fig. 1. (a) Top: Traditional one-stage RS-IUS architecture. 100% of the semantic information gap from sensory data to LC classes is filled up in one step. (b) Middle. Traditional two-stage noniterative GEOBIA design. 100% of the semantic information gap from sensory data to LC classes is filled up in the segment-based image classification second stage, in series with the subsymbolic inductive-data-learning image segmentation first stage. (c) Bottom. Novel three-stage hybrid RS-IUS design. Approximately, 50% of the semantic information gap from sensory data to LC classes is filled up in the automatic deductive preclassification first stage [80].

statistical preclassifier, implemented as a traditional plug-in (nonadaptive to input data) pixel-based maximum likelihood (ML) classifier. In this scenario, the ATCOR-SPECL [6]–[9] and SIAM [10]–[19] software products, to be assessed and compared in the Part 2 of this paper [20], are, to the best of these authors' knowledge, the first examples of prior knowledge-based decision-tree preclassifiers in operating mode eligible for use at the preattentive vision first stage of a hybrid RS-IUS architecture, see Fig. 1(c). Noteworthy, the hybrid RS-IUS architecture shown in Fig. 1(c) is alternative to both the two-stage hybrid RS-IUS architecture proposed by Shackelford and Davis [61], [62] and the GEOBIA/GEOOIA system architecture shown in Fig. 1(b). To summarize, whereas prior knowledge-based decision-tree classifiers have been traditionally employed in one-stage RS-IUSs [see Fig. 1(a)] or at the attentive vision second stage of two-stage hybrid

RS-IUSs, whose first stage consists of either a subsymbolic statistical system, like in GEOBIA/GEOOIA systems, see Fig. 1(b), or a semisymbolic plug-in statistical system, like in the Shackelford and Davis RS-IUS architecture [61], [62], the degree of novelty of the ATCOR-SPECL and SIAM prior knowledge-based preclassifiers is to provide a multistage hybrid RS-IUS architecture with an injection of prior knowledge right at the level of the preattentive vision first stage [10]–[19], see Fig. 1(c) [20]. Additional examples of hybrid inference systems for RS image classification are those proposed by Matsuyama *et al.* in [46], [47], as well as the popular Landsat-7 Enhanced Thematic Mapper $\mathrm{(ETM) + Automated}$ Cloud-Cover Assessment (ACCA) algorithm. In the ACCA algorithm, first, a per-pixel (context-independent) physical model-based decision rule set is applied to a radiometrically calibrated Landsat image to detect pixels considered as cloud candidates. Second, to remove small holes in cloud segments, a bottom-up (data-driven) context-sensitive aggregation and filling algorithm is applied in the (2-D) image domain to pixels considered as noncloud candidates at step one [63] (p. 1183).

C. Human and Computer Vision

In the words of Iqbal and Aggarwal: "frequently, no claim is made about the pertinence or adequacy of the digital models as embodied by computer algorithms to the proper model of human visual perception... This enigmatic situation arises because research and development in computer vision is often considered quite separate from research into the functioning of human vision. A fact that is generally ignored is that *biological vision is currently the only measure of the incompleteness of the current stage of computer vision, and illustrates that the problem is still open to solution*" [64].

According to this quote, human vision should be considered the gold standard (reference baseline) of the computer vision discipline, which incorporates RS image understanding as a special case. Unfortunately, the great majority of the RS community does not appear to consider biological vision as a reference baseline. In addition, relationships between the RS and computer vision communities appear weak too, the latter community considering the expertise of the former not very advanced, because traditional RS image understanding is pixel-based, where spatial (contextual) information is ignored. As a result of this lack of interdisciplinary communication, the RS community tends to underestimate the complexity of vision in general and RS image understanding in particular.

In the rest of this paper, including the experimental Part 2 [20], RS image understanding (classification, mapping) is conceived as a subset of computer vision, where human vision is adopted as a reference standard, to compare the ATCOR-SPECL and SIAM software products as alternative implementations of a prior knowledge-based preclassification first stage in a hybrid RS-IUS architecture [10]–[20] (refer to Section I). Hence, this paper complies with the aforementioned thesis by Iqbal and Aggarwal [64], but is in contrast with the majority of the RS literature, where links to computer vision and human vision disciplines are absent.

In this section, basic principles of human vision, which comprises a preattentive vision first stage and an attentive vision second stage, are briefly described [5], [46].

1) Goal of a (Biological or Artificial) Vision System: A (human or computer) visual system is a (biological or artificial) IUS suitable to provide plausible (multiple) symbolic description(s) of a 3-D scene, located in the (4-D) world-through-time, as it is observed by a (2-D) imaging sensor at a given acquisition time. The information gap between a subsymbolic (2-D) image and a symbolic (3-D) scene can be filled by conjectures that map subsymbolic image features (e.g., image-objects or, vice versa, image-contours) into symbolic classes of 4-D objects-through-time (4-D concepts-through-time, e.g., buildings and roads) belonging to the so-called preexisting (4-D) *world model* [46], [65]. A world model, also called *world ontology*, can be graphically represented as a semantic network consisting of: 1) classes of 4-D objects-through-time as nodes and 2) inter-concept relations as arcs between nodes, namely: (I) spatial relations, either topological (e.g., adjacency and inclusion) or nontopological (e.g., distance and in-between angle), (II) temporal relations and/or (III) nonspatiotemporal relations (e.g., part-of and subset-of) [18], [19], [46], [55], [66].

In terms of computational theory, the problem of image understanding (vision), from subsymbolic (2-D) imagery to symbolic description(s) of the (3-D) scene of the (4-D) world observed at a given time, belongs to the class of symbolic inductive data learning problems [24] (from sensory data to models, refer to Section II-B). As such, it is inherently ill-posed in the Hadamard sense [42] and, consequently, very difficult to solve, due to the combination of the two following *qualitative* and *quantitative information gaps* to be filled (refer to Section II-A) [18], [19], [46]: 1) The well-known *(semantic) information gap* between continuous subsymbolic sensory sensations and discrete symbolic (semantic, linguistic) persistent (stable) percepts (concepts), which has been thoroughly investigated in both philosophy and psychophysical studies of perception. In practice, "we are always seeing objects we have never seen before at the sensation level, while we perceive familiar objects everywhere at the perception level" [46]. 2) The intrinsic insufficiency of image features, namely, 0-D points, 1-D lines (e.g., contours) and 2-D polygons (image-objects), in the reconstruction of an observed (3-D) scene, due to *data dimensionality reduction* which causes, e.g., occlusion phenomena.

2) Processing Elements and Modular Structure of the Human Visual System: In mammals, a vision system accomplishes a preattentive vision first phase and an attentive vision second phase, summarized as follows.

1) Preattentive (low-level) vision extracts picture primitives based on general-purpose image processing criteria independent of the scene under analysis. It acts in parallel on the entire image as a rapid (< 50 ms) scanning system to detect variations in simple visual properties [67]–[69]. In the primary visual cortex (PVC, or area 17 of the visual cortex, or V1), single opponent and double opponent color cells are called Type I and Type II, respectively, by Wiesel and Hubel [72] (examples of Type I and Type II receptive fields can be found in [73]). Receptive fields that are spatially opponent, but not color

opponent, are called Type III [73]. Layers of PVC are vertically organized into blobs and interblob areas. The same single-opponent cells are thought to provide, in parallel, color contrast information to cells in the blobs, and achromatic contrast information to cells in the interblob regions. The visual cells heavily concentrated in cortical blobs are double-opponent cells. In the interblob areas, cortical cells belong to the hierarchy composed of simple- and complex-cell categories. A major difference between simple- and complex-cells is that the former are quasilinear while the latter exhibit a clear second-order squaring nonlinearity [98]. A regular sequence of hypercolumns is repeated over the surface of PVC, each hypercolumn occupying an area of about 1 mm^2. This repeating organization constitutes the modular structure of PVC, such that every axis of orientation, whose gradations of orientation are around 10° [67] to 15° [70], [71], is represented for every retinal position at at least four spatial scales of analysis [99]. In each hypercolumn, there are end-stopped cells, in addition to simple- and complex-cells [100]. While simple- and complex-cells are thought to accomplish line and edge extraction, end-stopped cells respond to image singularities, such as line/edge crossings, vertices of image-objects, and end-points of line segments [101].

2) Attentive (high-level) vision operates as a careful scanning system employing a focus of attention mechanism based on end-stopped cells [100], [101]. Scene subsets, corresponding to a narrow aperture of attention, are observed in sequence and each step is examined quickly (20–80 ms) [67]–[69].

It is worth noting that human achromatic vision is nearly as effective as human chromatic vision in detecting forms and accomplishing image interpretation. On an *a posteriori* basis, this observation has two important implications. First, in the real 4-D world-through-time, color information of 4-D objects (e.g., cars and trees) is dominated by their 4-D spatiotemporal information, as properly stated by Adams *et al.* [26]. Second, the same consideration holds for a (2-D) image representation of the (4-D) world-through-time, where 2-D spatial (contextual) information dominates color information. To cope with the dominant 2-D spatial information in a (2-D) image, the human visual system employs modular arrays of multiscale 2-D local filters capable of providing a topology-preserving mapping of a (2-D) image [67]–[71], [74].

3) When Does Vision Go Symbolic? Inference Mechanisms in Human Vision: In the literature of psychophysics, according to Vecera and Farah, *preattentive image segmentation is an interactive (hybrid) inference process "in which top-down knowledge (e.g., familiarity) partly guides lower level processing"* ([75]; p. 1294). That is to say, *human vision is a symbolic hybrid (combined deductive and inductive) inference system where (symbolic) prior knowledge is injected into the sensory data interpretation process starting from the preattentive vision first stage* [18], [19].

In the computer vision literature, according to Marr "*(human) vision goes symbolic almost immediately, right at the level of (second-order derivative's) zero-crossing (raw primal sketch)... without loss of information*" ([5]; p. 343), which is consistent with the aforementioned quote by Vecera and Farah [75]. Unfortunately, in [5], the computer vision system proposed by Marr is unable to satisfy either one of the two aforementioned vision system requirements inspired by human vision. In particular, the Marr preattentive vision first stage is subsymbolic. It is split into a subsymbolic raw primal sketch and a subsymbolic full primal sketch, where: (I) the raw primal sketch consists of a hierarchy of subsymbolic primitives, namely, multiscale zero-crossings ([5]; pp. 54–59), followed by zero-crossing segments ([5]; p. 60) and level 1 image-tokens, comprising blobs (closed contours), edges, bars and discontinuities (terminations) ([5]; pp. 70–73), and (II) a full primal sketch, equivalent to perceptual grouping [75]–[77], where level 2 boundaries (e.g., texture boundaries) are detected between groups of tokens ([5]; pp. 53, 91–95). Marr never provided implementation details of his proposed subsymbolic raw primal sketch or subsymbolic full primal sketch. This apparent contradiction between Marr's computer vision system design (computational theory) specifications and his own implementation is not at all surprising. It accounts in general for the customary distinction between a model and the algorithm used to identify it [18].

4) Possible Relationships Between a Human Vision System and the ATCOR-SPECL and SIAM Prior Knowledge-Based Preclassifiers: Possible relationships between a human vision system, as it is described in Sections II-C1–II-C3, and the ATCOR-SPECL and SIAM prior knowledge-based preclassifiers, to be investigated in the Part 2 of this paper as alternative implementations of a preattentive vision first stage in a hybrid RS-IUS architecture [20], are highlighted as follows.

1) At the abstraction level of computational theory (system design), the hybrid RS-IUS architecture proposed in this paper is consistent with a human vision system conceived as a symbolic hybrid inference system where symbolic prior knowledge is injected right at the preattentive vision first stage (see Section II-C3).

2) In (2-D) images of the (4-D) world-through-time, 2-D spatial (contextual) information dominates color information (see Section II-C2). In traditional pixel-based RS-IUSs, the input data set is a 1-D sequence of pixel-specific data vectors where 2-D space (contextual) information is ignored. A pixel-based RS-IUS can perform accurately without 2-D spatial information in the image domain if and only if the image spatial resolution and time resolution are adequate to discriminate the target phenomenon under investigation based on (context-insensitive) color-through-time properties exclusively. It means that, to be considered useful, the application-independent ATCOR-SPECL and SIAM prior knowledge-based preclassifiers, which are pixel-based (context-insensitive) and eligible for use with any single-date RS imagery independent of its spatial resolution, must be considered as simple building blocks in a multistage RS-IUS architecture, i.e., they cannot be considered as standalone systems. In fact, their first-stage pixel-based (color-driven) preattentive image analysis must be followed by an attentive vision second stage, capable of (2-D) spatial analysis plus 1-D temporal

analysis of image data conditioned (driven, stratified) by first-stage spectral categories, equivalent to conventional color names to be community agreed upon [102], [103]. In terms of filling the information gap from sensory data to LC maps (refer to Section II-C1), the ATCOR-SPECL and SIAM prior knowledge-based preclassifiers map subsymbolic sensory data into semisymbolic spectral categories (refer to the further Section IV-B) based on single-date pixel-based MS (color) properties (spectral signatures) exclusively. The remaining information gap from semisymbolic spectral categories to LC classes must be filled by the RS-IUS' attentive vision second stage based on stratified spatiotemporal information.

We can conclude that, if compared with a human visual system, the degree of compatibility of the ATCOR-SPECL and SIAM prior knowledge-based preclassifiers, employed in support of the preattentive vision first stage of a hybrid RS-IUS architecture, is inferior to the degree of biological plausibility of an airplane compared to a bird. That said, from an engineering standpoint, the ATCOR-SPECL and SIAM deductive preclassifiers provide a realistic and feasible contribution to the development of automatic hierarchical RS-IUSs in operating mode, where a preattentional first-stage prior knowledge-based discretization of a continuous color space may be employed to better condition for numerical treatment an inherently difficult-to-solve second-stage attentive vision spatio-temporal analysis.

D. EO Big Data: Challenges and Opportunities

According to Section I, the secondary objective of this paper is to contribute to the development of a new generation of operational hybrid RS-IUSs capable of transforming large-scale multisensor multiresolution EO image databases into information products, in compliance with the QA4EO guidelines. The magnitude of EO data collected since the early 1970 s by a variety of spaceborne/airborne and *in situ* sensory data sources, at varying spatial extents and multiple spatial, temporal and spectral resolutions, is so phenomenal to be identified, by the present authors, as EO big data, in line with the terminology of IT.

In IT, the popular term "big data" identifies "a collection of data sets so large and complex that it becomes difficult to process using on-hand database management tools or traditional data processing applications. The challenges include capture, storage, search, sharing, analysis, and visualization" [78]. Among big data challenges, interpretation of observational data, i.e., the transformation of sensory data into information/knowledge products, has been historically investigated by both philosophical hermeneutics [21], [22] (refer to Section II-A) and psychophysical studies of perception [46] (refer to Section II-C).

According to the present authors, "big data" is a synonym of "central limit theorem." In statistics, the well-known central limit theorem states that [29], given certain conditions (typically random variables must be identically distributed), the sum (mean) of a sufficiently large number of independent random variables, each with a well-defined mean and well-defined variance (for example, one random variable is an LC class-specific distribution of pixel values in a RS image), tends to form a Gaussian distribution, where no "meaningful" or "natural" hidden data entities, clusters or (sub)structures can be identified [18], [19]. As a consequence of the central limit theorem, "big data" distributions are Gaussian-like, hence meaningful cluster/substructure detection in "big data" is inherently ill-conditioned in the Hadamard sense (refer to Section II-B). In other words, in "big data" sets, traditional inductive supervised or unsupervised data learning is extremely difficult or impossible to accomplish (refer to Section II-B).

These general considerations, driven from common knowledge in IT, may explain why, to date, EO big data assets are underemployed by the RS community. For example, the European Space Agency (ESA) estimates as 10% or less the percentage of RS images ever downloaded (which does not mean ever used) by stakeholders from its EO databases [18], [19]. It may mean that the RS discipline is still incapable of filling up the information gap from RS data to knowledge/information products (refer to Section II-C). To fill this information gap, data interpretation (cognitive) processes (related to the concept of equivocal "information-as-(an intepretation)process") dominate, i.e., are more difficult to solve than data transformation (e.g., data enhancement, data preprocessing) tasks (related to the concept of unequivocal "information-as-thing," refer to Section II-A). Typically, RS scientists and practitioners overlook their cognitive inadequacy to derive "operational, comprehensive, and timely knowledge/information products" from sensory data [1]–[3] by asking for more data of better quality, which actually makes their cognitive lack even worse. In practice, by overestimating its data interpretation capability the RS community is outpaced by the ever-increasing rate of collection of EO data of enhanced quality and quantity [10]–[19] (also refer to the further Section III).

To recapitulate, in agreement with common knowledge in IT, EO big data assets represent a huge opportunity/challenge for the RS interdisciplinary science. To be transformed into knowledge/information products in compliance with the QA4EO guidelines [1]–[3], EO big data require the development of a novel generation of hybrid inference systems in operating mode, capable of outperforming traditional inductive or deductive inference systems, whose limitations are well known (refer to Section II-B). As a realistic contribution to this challenge, this paper provides a theoretical and experimental assessment of the ATCOR-SPECL and SIAM prior knowledge-based preclassification software products in operating mode.

E. Probability and Nonprobability Sampling of a Geospatial Population

This paper requires that sample QIs, estimated from the ATCOR-SPECL and SIAM deductive preclassification maps, must be statistically valid (consistent), refer to Section I. By definition, an information map (e.g., a thematic map) is a reduced representation of a target geospatial population. To provide a statistically valid estimation of QIs from an information map representing a geospatial population [82], [83] (refer to

Section I), the following definitions of probability and nonprobability sampling protocol are required.

1) By definition, probability sampling must satisfy three necessary not sufficient conditions to deliver statistically valid sample estimates, i.e., sample estimates provided with the necessary probability foundation to permit generalization from the sample data set to the whole target geospatial population being sampled [82], [83]. 1) All inclusion probabilities must be greater than zero in the target geospatial population to be sampled. If some sampling units have an inclusion probability of zero, then the accuracy assessment does not represent the entire target region depicted in the map to be assessed and the results cannot be deemed statistically consistent. 2) The inclusion probabilities must be: a) knowable for nonsampled units and b) known for those units selected in the sample: since the inclusion probability determines the weight attached to each sampling unit in the accuracy estimation formulas, if the inclusion probabilities are unknown, so are the estimation weights. Probability sampling methods can be split into equal or variable (unequal) probability sampling methods. Unequal inclusion probabilities create no difficulties as long as they are known for sampled units and accounted for in the estimation formulas, but equal probability designs are advantageous in that they allow for simpler analysis. For example, an area sampling protocol selects polygons into the sample with an inclusion probability monotonically increasing with the polygon area [82], [83]. Noteworthy, no probability sampling is required to assess the degree of uncertainty in sample estimates [5].
2) Nonprobability sampling methods do not satisfy the requirements of probability sampling methods listed in this section above. According to the existing literature [82]: "unfortunately, examples of nonprobability sampling are common in accuracy assessment applications. Selecting reference locations by purposeful, convenient, or haphazard procedures does not allow the sampling design to determine the inclusion probabilities for each sampling unit. Such designs, therefore, are not probability samples. Purposefully, selecting training data for a supervised classification is a good example of a nonprobability sample. Such samples are acceptable for developing a land cover classification map, but often have limited use for accuracy assessment because the necessary probability foundation to permit generalization from the sample data to accuracy of the full population is lacking." To recapitulate, "it is possible to obtain useful information from nonprobability samples, but the limitations of such data should be recognized" [82]. For example, nonprobability sampling allows to assess the degree of uncertainty in sample estimates.
3) A protocol, defined as a sorted set of guidelines for good practice [3], encompasses a *structural knowledge* and a *procedural knowledge*, like in decision trees [55]. Structural knowledge is related to the content of the rule set while procedural knowledge is related to the order of presentation of rules. The definition of international protocols for best practices, such as the QA4EO guidelines [2], together with standardization, have been major challenges for the RS community [2], [3].

Unfortunately, in the RS literature there is a lack of probability sampling protocols adopted for the validation of RS data-derived products in compliance with the principles of statistics and the QA4EO guidelines. As a negative example of nonprobability sampling for map quality assessment not to be imitated, refer to [41].

A probability sampling protocol for thematic and spatial quality assessments of classification maps generated from EO images is proposed in [80] and adapted in Part 2 of this paper [20].

F. QIO of an RS-IUS

The test phase of a software product, which encompasses a QI selection stage, can be so relevant to absorb up to 50% of a project budget [93]. In this section, a possible list of mutually uncorrelated metrological/statistically-based QIOs is proposed and recommended for use by the Part 2 of this paper, to accomplish the experimental assessment and comparison of the ATCOR-SPECL and SIAM software products in operating mode [20].

Often forgotten in practice, the noninjective property of any metrological/statistically-based QI states that it is always possible to find two different instances of the same target phenomenon capable of generating the same QI value. For example, two different classification maps may provide the same map's overall accuracy value. This is tantamount to saying that no universal QI can exist [10], [19], which is in contrast with a significant segment of the existing literature, e.g., see [79] and [94]. Rather, a target-specific set of complementary statistically independent QIs must be selected and agreed upon by the scientific community.

To cope with EO big data challenges (refer to Section II-D), this paper provides an assessment of operational RS-IUSs in compliance with the principles of statistics, the QA4EO guidelines [2] and the GEO-CEOS land product accuracy validation criteria [3] (refer to Section I). These work requirements mean that the quality assessment of an RS-IUS should rely on a complete set of complementary metrological/statistically-based QIOs that are statistically independent, valid and significant. To be considered statistically significant, QIOs must be provided with a degree of uncertainty in measurement (refer to Section I). To be statistically valid (consistent), QIOs must be estimated from probability sampling of EO big data (refer to Section II-E).

Selected from the existing literature, a possible list of QIOs of an information processing system in operating mode is proposed as follows, to be community-agreed upon [10]–[19]. 1) Degree of automation (ease-of-use), monotonically decreasing with the number of system free-parameters to be user-defined based on heuristics. 2) Effectiveness, e.g., thematic accuracy and spatial accuracy of classification and segmentation maps generated from EO images [80]. 3) Efficiency, e.g., inversely related to computation time and memory occupation. 4) Robustness to changes in input parameters, if any free-parameter exists. 5) Robustness to changes in input data acquired across time, space and sensors. For example, refer to the CEOS land product accuracy validation stages 1–4 in [3], [4]. 6) Scalability, to cope with changes in input data specifications, sensors and user's requirements. 7) Timeliness, defined as the time between data acquisition and data-derived high-level product generation. For example, user interactions, such as those required to collect reference samples for training a supervised data learning system, increase timeliness [81]. 8) Costs, monotonically increasing with computer power and manpower.

To be termed operational, an information processing system must score high in every QIO of a set of community-agreed independent QIOs, e.g., refer to points 1) to 8) in the previous paragraph.

Unfortunately, experiments presented in large portions of the RS literature are affected by the following methodological drawbacks. 1) The sole mapping accuracy is selected from the possible set of mutually independent QIOs eligible for parameterizing RS-IUSs for assessment and comparison purposes. 2) Statistical estimates of the mapping accuracy are not provided with a degree of uncertainty in measurement, i.e., they have no statistical significance. 3) Statistical estimates of the mapping accuracy are not collected by means of a probability sampling strategy, hence they lack statistical consistency (refer to Section II-E). 4) Alternative RS data mapping solutions are tested exclusively in toy problems, defined in this paper as test data mapping problems featuring a small spatial scale (e.g., local scale) and/or a coarse semantic granularity, such that these test cases do not reflect the complexity of the existing "EO big data" archives (refer to Section II-D) that must be dealt with to comply with the QA4EO requirements [2] (refer to Section I). As a consequence of these experimental limitations, many RS-IUS implementations tested in the RS literature feature the following drawbacks. (I) A mapping accuracy which remains unknown in statistical terms and/or is unable to generalize from a sample data set to the whole target geospatial population being sampled. (II) A robustness to changes in the input data set which is unknown or appears questionable. (III) A scalability to real-world RS data applications at large (e.g., continental and global) spatial scale and fine semantic granularity which is unknown or appears questionable.

The conclusion of this section is that, in real-world RS data applications, different from toy problems at small spatial scale and/or coarse semantic granularity, published RS-IUSs are likely to score poorly in operating mode, because at least one of their OQI values is expected to score low.

III. Problem Recognition and Opportunity Identification: Compliance of Existing RS-IUS Commercial Software Products With the QA4EO Key Principles And Calibration/Validation (*Cal/Val*) Requirements

Adopted by the ongoing GEOSS implementation plan for years 2005–2015 [1], the international GEO-CEOS QA4EO recommendations promote the development of "operational, comprehensive, and timely knowledge/information products" from a variety of satellite, airborne, and *in situ* sensory data sources [2] (refer to Section I). To guarantee "the provision of and access to the Right Information, in the Right Format, at the Right Time, to the Right People, to Make the Right Decisions," the QA4EO guidelines require the successful implementation of two necessary and sufficient key principles [2]: (I) *Accessibility/Availability* and (II) *Suitability/Reliability* of RS data and data-derived knowledge/information products (refer to Section II-A). To accomplish these system requirements the GEO identified the need to develop a GEOSS data quality assurance strategy where calibration and validation (*Cal/Val*) activities become critical to data quality assurance and, thus, to data usability. According to the QA4EO guidelines [2], [3], the following *Cal/Val* activities are required.

1) *An appropriate coordinated program of calibration activities throughout all stages of a spaceborne mission, from sensor building to end-of-life, is considered mandatory to ensure the harmonization and interoperability of multisource multitemporal RS data* [2]. By definition, radiometric calibration is the transformation of dimensionless digital numbers (DNs) into a community-agreed physical unit of radiometric measure, e.g., TOA radiance (TOARD), TOA reflectance (TOARF), and spectral reflectance (SURF).
2) To satisfy validation requirements (e.g., accuracy validation [3]), *observational data and data-derived products generated in each step of a satellite-based information processing workflow must have associated with them a set of independent, quantifiable, metrological/statistically-based QIs, featuring a degree of uncertainty in measurement at a known degree of statistical significance*, to comply with the general principles of statistics and provide a documented traceability of the propagation of errors through the information processing chain in comparison with established "community-agreed reference standards" [2] (refer to Section II-F).

It is an indisputable fact that, to date, almost ten years from the launch of the GEOSS initiative, the RS community has been more successful in pursuing the first rather than the second GEOSS key principle. For example, in line with the GEOSS requirement of Accessibility/Availability of RS data and data-derived products, the U.S. 2008 free Landsat data policy has opened a new era of exploitation of the more than three million scenes stored in the U.S. Landsat archive [84]. On the other hand, the ever-increasing rate of collection of EO data of enhanced spatial, spectral and temporal quality outpaces the current ability of the RS discipline to transform EO big data assets into knowledge/information products (refer to Section II-D). This means that the GEOSS requirement of Suitability/Reliability of sensory data and data-derived products can still be considered far from being accomplished by the RS community.

To explain their different degrees of success, the first and second GEOSS key principles are analyzed at different levels of abstraction. At the abstraction level of knowledge/information representation, according to philosophical hermeneutics [21],

[22], the first GEOSS key issue is quantitative (unequivocal) and related to the Shannon concept of "information-as-thing" irrespective of its meaning [25]. As such, it is easier to deal with than the second GEOSS principle, which is qualitative (equivocal), since the latter has to deal with the meaning (interpretation, understanding) of sensory data and is related to the concept of "information-as-(an interpretation) process" [21], [22] (refer to Section II-A).

At the abstraction level of RS-IUS design, the second GEOSS key principle remains difficult to cope with also because *Cal/Val* activities are often neglected or ignored in the RS common practice. On theory, the RS community regards as common knowledge that *"the prerequisite for physically based, quantitative analysis of airborne and satellite sensor measurements in the optical domain is their calibration to spectral radiance"* ([95], p. 29). Moreover, according to related works [10]–[19], radiometric calibration is a necessary not sufficient condition for automatic interpretation of (for physical model-based inference from) EO imagery, refer to Section II-B. On the other hand, RS scientists, practitioners and institutions tend to overlook *Cal/Val* activities as necessary not sufficient preconditions for the harmonization of large-scale multitemporal multisensor EO datasets. For example, the European Commission Image 2000 product is a noncalibrated multisensor MS image mosaic at European scale, whose scientific usability for quantitative variable estimation is questionable or null [96]. To recover from this lack, the European Commission Image 2006 program includes radiometric calibration of multisensor MS images at European scale in its project requirements specification. However, in the Image 2006 project, no RS data-derived product validation policy is enforced [97].

To explain why radiometric calibration is neglected in the RS common practice, let us investigate the degree of compliance of RS-IUS commercial software products with the QA4EO key principles and *Cal/Val* requirements. Starting from the RS-IUS architectures proposed in Fig. 1, consider the: 1) two- or three-stage Trimble eCognition Developer [60], 2) one- or two-stage Pixel- and Segment-based versions of the Environment for Visualizing Images (ENVI) by ITT VIS [85], 3) one- or two-stage IDRISI Taiga, 4) one-stage ESRI ArcGIS, 5) ATCOR-2/3/4 [6]–[8], 6) one-stage PCI Geomatica (with an optional ATCOR for atmospheric correction), and 7) one- or two-stage ERDAS IMAGINE Objective (with an optional ATCOR for atmospheric correction). These commercial software packages for RS image processing/ understanding consist of large suites of options to choose from [18], [56]–[59]. Frequently considered overwhelming by nonexpert users, these large software suites allow selectable algorithms to be chosen, supervised, and combined by a user, based on heuristics, to form a user- and application-specific information processing workflow. Among these wide sets of selectable algorithms, several options may appear not particularly relevant, or be difficult to use (because they require lots of user interactions to run) or omit steps considered critical in a standard RS data processing chain (like those promoted by the QA4EO recommendations [2]). In practice, to favor flexibility considered necessary to develop customized solutions, these software suites promote an approach to RS image analysis closer to art, namely, empirical, qualitative and nonreproducible, than science, which is rigorous, quantitative and reproducible. For example, the large majority of selectable algorithms implemented in the RS-IUS commercial software products listed above, with the sole exception of the physical model-based ATCOR-2/3/4 toolbox [6]–[8], does not consider radiometric calibration as mandatory. This relaxed input data constraint means that, in these commercial software products, the large majority of selectable algorithms consist of statistical systems, hence the remaining small minority comprises physical models. Due to their inherent ill-posedness in the Hadamard sense [42], statistical systems are typically semiautomatic and site-specific [18], [45] (refer to Section II-B). Although statistical systems do not require as input observational data provided with a physical meaning, they may benefit from radiometric calibration in terms of robustness to changes in the input data set (refer to Section II-B). For example, in the ENVI commercial software toolbox [85], an atmospheric correction tool, called Fast Line-of-sight Atmospheric Analysis of Spectral Hypercubes (FLAASH), is available as an optional RS image preprocessing stage. As another example, in the PCI Geomatica and ERDAS RS data preprocessing workflows, a physical model-based ATCOR module can be optionally installed, etc.

The first conclusion about the RS-IUS commercial software products listed above is the following. In line with common knowledge in the machine learning community [24], since statistical model-based systems are inherently poorly-conditioned, semiautomatic and site-specific and require prior knowledge in addition to data to become better posed for numerical treatment (refer to Section II-B), then statistical systems available for selection in RS-IUS commercial software products, where they typically outnumber physical model-based options, are expected to be, *per se*, unable to cope with the well-known challenges of EO big data (refer to Section II-D). To become more successful, these statistical systems must be combined with physical models, to form hybrid inference systems capable of outperforming their individual components (refer to Section II-B). This consideration holds because at least one or more QIOs (e.g., timeliness, scalability, and robustness to changes in the input data set, refer to Section II-F) of any inductive data learning system, either supervised or unsupervised, whether or not it adopts an RS data radiometric calibration preprocessing stage in compliance with the QA4EO guidelines (refer to Section III), are expected to score low in real-world RS data mapping applications (refer to Section II-B), where EO big data assets (refer to Section II-D), different from unrealistic toy problems at small spatial scale and/or coarse semantic granularity (refer to Section II-F), are to be mapped.

In addition, RS-IUS commercial software products, such as those listed above, appear affected by a lack of selectable physical model-based inference systems, considered necessary to support, with prior knowledge in addition to data (in accordance with well-known principles of inductive inference, clearly stated by Mulier and Cherkassky [24], refer to Section II-B), the large majority of selectable options, consisting of statistical systems. This second conclusion about the RS-IUS commercial software products listed above is driven from the sole physical model found in this list, the ATCOR [6]–[8].

The core of the ATCOR consists of a radiative transfer model which is inverted to calculate as output directional surface reflectance (SURF) values starting from at-sensor (top-of-atmosphere, TOA) radiance (TOARD) values [9]. In the standard ATCOR implementation, the influence of surface type-specific bidirectional reflectance distribution function (BRDF) effects is not modeled. In the words of the ATCOR's authors [9], "ideally, an atmospheric and radiometric correction routine would result in BRDFs for all observed targets, as the BRDF is the unambiguous radiometric property of the Earth's surface. Unfortunately, imaging spectrometers rarely provide sufficient information to produce reliable BRDFs as most instruments acquire data for a single view geometry. Thus, a quantity not depending on the view geometry is of interest. The spectral albedo, i.e., the bihemispherical reflectance (BHR), is a value which is well suited for an unbiased view of the Earth's surface." In recent years, an "augmented" ATCOR implementation, sketched in Fig. 2, has been tested to retrieve spectral albedo in series with surface reflectance values starting from dimensionless DNs [9]. A peculiar aspect of this augmented ATCOR workflow, suitable for continuous variable estimation from RS data, is that categorical variables are generated as intermediate products by preliminary classification modules at several hierarchical stages (refer to Section II-A). In Fig. 2, data processing blocks identified as "preclassification" and "quantitative classification" are suitable for mapping semantic concepts from data, such as "clouds," "water," "vegetation," and "haze." Once estimated from sensory data, these categorical variables are further employed as input to processing modules capable of continuous (e.g., bio-physical) variable estimation (refer to Section II-B). That is to say, in the augmented ATCOR workflow shown in Fig. 2, the inherently poorly-conditioned inductive inference problem of continuous variable estimation from sensory data is accomplished on a symbolic stratified (driven-by-knowledge) basis to become better conditioned for numerical treatment (refer to Section II-B). In practice, the complete atmospheric correction and radiometric normalization scheme shown in Fig. 2 provides an additional source of experimental evidence supporting the recent conjecture, proposed in the RS literature [15], [80], that *categorical variables (e.g., LC and LCC maps) and continuous variables (e.g., spectral albedo, LAI and green biomass), conceived as two sides of the same coin, should be estimated from RS images alternately and iteratively, starting from a categorical variable estimation first stage* (refer to Section I). *Intuitively, MS image preclassification is preliminary to continuous variable estimation, which includes atmospheric correction, because the former task is "easier" to accomplish than the latter*. In fact, an expert photointepreter can successfully interpret (classify) an RS image irrespective of whether this image has been provided with a physical unit of radiometric measure through radiometric calibration. On the other hand, the RS literature clearly acknowledges that no spectral index (e.g., the normalized difference vegetation index, NDVI) should ever be computed as a quantitative proxy of a continuous biophysical variable (e.g., a LAI value), if no radiometric calibration has taken place, yet [45].

To summarize, capable of alternating categorical and continuous variable estimation from sensory data, the surface albedo estimation workflow shown in Fig. 2, based on an inverted radiative transfer model, is provided with a relevant degree of novelty in comparison with standard radiative transfer software products, like the Second Simulation of the Satellite Signal in the Solar Spectrum (6S) [86]. For example, in the 6S software tool, the land cover class-specific BRDF effects correction of RS imagery relies on ancillary thematic information, i.e., the 6S software product is per se unable to extract from the input RS image the surface types (e.g., ocean surface, vegetation and bare soil [86]) required as input to run the driven-by-knowledge BRDF correction phase.

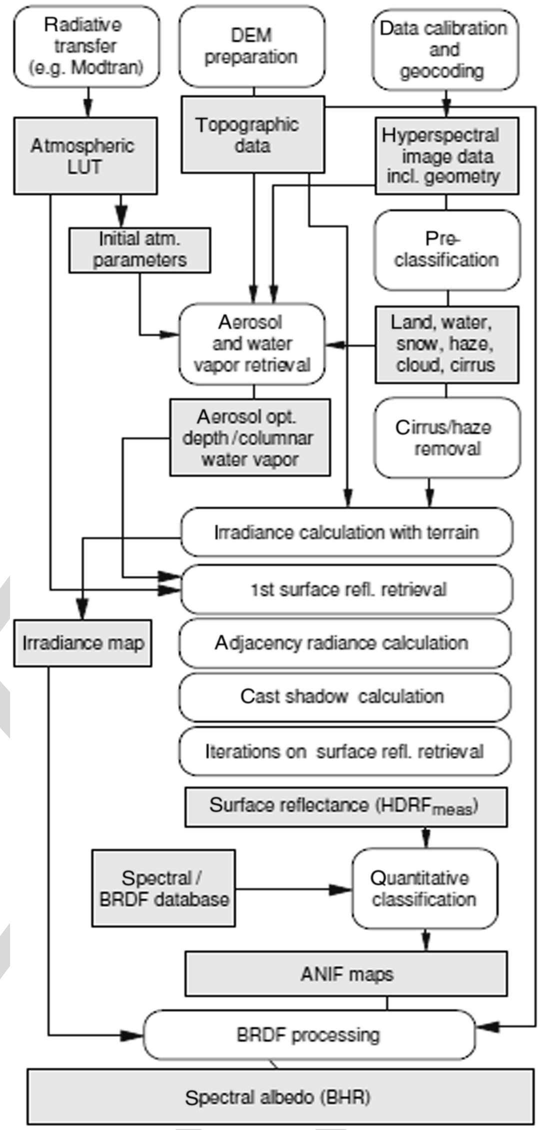

Fig. 2. A complete ("augmented") physical model-based system for RS data normalization combines a standard ATCOR workflow [6]–[9] with a novel bidirectional reflectance distribution function (BRDF) effect correction. Processing blocks are represented as circles and output products as rectangles. This workflow estimates categorical and continuous variables from sensory data alternately, starting from a prior knowledge-based pre-classification first stage, such as SPECL. Same as in [9], courtesy of Daniel Schläpfer, ReSe Applications Schläpfer.

This section concludes that, eligible for use as the physical model-based "preclassification" block in Fig. 2, *the ATCOR-SPECL and SIAM prior knowledge-based preclassifiers feature* *a wide application domain, encompassing not only categorical variable estimation from EO data* (as it is logical to expect from a preclassification system), *but also continuous variable estimation from EO data, in compliance with the Cal/Val activities considered mandatory by the QA4EO guidelines for both RS data preprocessing (data enhancement) and RS data processing (data understanding) phases* [2]. In other words, the ATCOR-SPECL and SIAM deductive preclassifiers appear as viable tools to accomplish not only automatic mapping of real-world EO big data sets (refer to Section II-D), in compliance with the QA4EO guidelines and the objectives of this paper (refer to Section I), but also RS image enhancement, as shown in Fig. 2. Existing examples of the SIAM applied to RS image preprocessing problems include stratified topographic correction [15], stratified atmospheric correction [6]–[8], stratified image mosaicking, stratified image co-registration, etc. [10]–[19] (refer to the further Section IV-A).

IV. Comparison of the ATCOR-SPECL and SIAM Software Products at the Four Levels of Understanding of an Information Processing System

Starting from the interdisciplinary nomenclature introduced in Section II, differences and similarities between the ATCOR-SPECL and SIAM software products can be investigated at the four levels of abstraction of an RS-IUS [5], [16], [18], [30], [87], namely: 1) computational theory (system architecture), 2) information/knowledge representation, 3) algorithms, and 4) implementation. Among these four levels of analysis, the first two are considered of fundamental importance for the success of any information processing system in operating mode (refer to Section I). In the words of Sonka *et al.*, "*the linchpin of success (of an information processing system) is addressing the (computational) theory (and information/knowledge representation [87]) rather than algorithms or implementation*" ([30]; p. 376).

A. Computational Theory

In Section I, the ATCOR-SPECL and SIAM software products are introduced as two alternative prior knowledge-based color space discretizers capable of providing a hybrid RS-IUS architecture with an injection of prior spectral knowledge, equivalent to color naming, right at the preattentive vision first stage, in compliance with human vision (refer to Section II-C). Common features of the two deductive image mapping systems are the following. 1) As physical models, they require as input a MS image provided with a physical unit of measure, namely, a MS image radiometrically calibrated into TOARF or SURF or surface albedo values (refer to Sections II-B and III). 2) They are context-insensitive, i.e., pixel-based, because color is the sole (0-D) pixel-specific information in a (2-D) image. All remaining visual properties are context-sensitive, e.g., texture [73], shape of image-polygons, and inter-object spatial relations [10]–[19], [46], [47], [61], [62]. 3) They are static, i.e., nonadaptive to input data, 4) one-pass, i.e., noniterative, 5) syntactic, i.e., rule-based [30], 6) semisymbolic, i.e., eligible for mapping a MS image into a discrete and finite set (legend) of spectral-based semiconcepts (refer to Section I), and 7) "fully automatic," because deductive inference systems require neither user-defined parameters nor training data sample to run [88] (refer to Section I).

Since they share the aforementioned list of system specifications, then the ATCOR-SPECL and SIAM systems can be used interchangeably in a hybrid RS-IUS workflow, such as those shown in Fig. 1(c) or 2. Although interchangeable, the ATCOR-SPECL and SIAM prior knowledge-based preclassifiers are not expected to perform the same, since their decision-tree design and implementation are completely different, in terms of both structural and procedural knowledge (refer to Section II-E).

A novel three-stage hybrid RS-IUS architecture, shown in Fig. 1(c), whose preattentive vision first stage employs a prior knowledge-based preclassifier provided with feedback loops [10]–[19], is described as follows.

1) An EO image preprocessing stage zero, suitable for MS image enhancement, including a mandatory MS image radiometric calibration of DNs into TOARF values, in compliance with the QA4EO guidelines. Although SURF values, considered as a special case of TOARF values in very clear sky conditions and flat terrain conditions [12], [80], [89], i.e., $\mathrm{TOARF} \supseteq \mathrm{SURF}$, such that $\mathrm{TOARF} \approx \mathrm{SURF} + \mathrm{atmospheric}$ "noise," are allowed as input, they are not mandatory, i.e., atmospheric correction is not considered a MS image preprocessing requirement.
2) A physical model-based symbolic context-insensitive (pixel-based) preattentive vision first stage, like the ATCOR-SPECL or the SIAM prior knowledge-based preclassifier. An injection of prior knowledge in the preattentive vision first stage makes the inherently poorly-conditioned EO image interpretation problem better posed for numerical treatment (refer to Section II-B), in agreement with the Marr intuition that vision goes symbolic right at the level of the raw primal sketch [5] (refer to Section II-C).
3) A second-stage battery of attentive vision context-sensitive stratified (driven-by-knowledge) application-, sensor- and LC/LCC class-specific feature extractors (e.g., multiscale texture is investigated exclusively in the image portion masked by the first-stage spectral category "vegetation," in order to split spectral type "vegetation" into two LC classes, namely, low-texture "grassland" and high-texture "forest" [61], [62]) and one-class LC/LCC classification modules (e.g., if a first-stage spectral category mask is "vegetation" and the second-stage "vegetation" masked data feature extractor is "high texture," then "forest").
4) A feedback mechanism between the preattentive vision first stage, the attentive vision second stage and the RS image preprocessing stage zero. Existing examples of these feedback loops are stratified topographic correction [15], stratified atmospheric correction [6]–[8], stratified image mosaicking, stratified image co-registration, and cloud/cloud-shadow masking [10]–[19].

This novel hybrid RS-IUS design [see Fig. 1(c)] is alternative to the two-stage hybrid RS-IUS architecture proposed by Shackelford and Davis [61], [62], whose first stage is a nonadaptive statistical classifier, namely, a plug-in parametric ML classifier (refer to Section II-B), and to state-of-the-art two-stage noniterative GEOBIA system [see Fig. 1(b)] and three-stage iterative GEOOIA system architectures [18], [19] (refer to Section II-B), where: 1) the preattentive vision first stage consists of an unlabeled data learning algorithm for image segmentation [32]–[34], [55]–[60], which is inherently poorly-posed [24] and is, therefore, semiautomatic and site-specific [45]; and 2) prior knowledge, if any, is injected exclusively at the attentive vision second stage, if and only if this second stage is implemented as a static image-object-based decision-tree classifier. If no prior knowledge is employed at the GEOBIA/GEOOIA attentive vision second stage, because it is implemented as an inductive data learning classifier (e.g., an artificial neural network classifier, a support vector machine classifier [41], a nearest-neighbor classifier, an adaptive decision-tree classifier, and a radial basis function network for classification [24], [29]), then the GEOBIA/GEOOIA system implementation is fully inductive at both first and second stages, which means that the GEOBIA/GEOOIA system, due to its inherent ill-posedness, is semiautomatic and site-specific in common practice (refer to Section II-B). This line of reasoning justifies the low productivity of many GEOBIA/GEOOIA systems increasingly observed in the existing literature [56], [57], which makes them inadequate to cope with large-scale RS image databases.

B. Information/Knowledge Representation

The ATCOR-SPECL and SIAM software products are compared in terms of: 1) input MS data requirements and 2) output preclassification map's legend.

1) Input MS Data Requirements Specification: The physical model-based ATCOR-SPECL and SIAM prior knowledge-based preclassifiers require as input MS images radiometrically calibrated into a physical unit of radiometric measure (refer to Section II-B), in compliance with the *Cal/Val* requirements of the QA4EO guidelines [2] (refer to Section III). In particular, SIAM requires as input a MS image radiometrically calibrated into TOARF or SURF or surface albedo values, where SURF is a special case of TOARF in very clear sky conditions and flat terrain conditions [12], [80], [89], i.e., $\mathrm{TOARF} \supseteq \mathrm{SURF}$, such that $\mathrm{TOARF} \approx \mathrm{SURF} + \mathrm{atmospheric}$ "noise." It means that an LC class-specific family of spectral signatures in TOARF values forms a buffer area (envelope) which includes, as a special case, the family of "ideal" (atmospheric noiseless) spectral signatures in SURF values for that same LC class, see Fig. 3.

In practice, SIAM is capable of recognizing surface types in RS images by "looking through" atmospheric effects, like the presence of haze and thin clouds [10]–[19]. This "look-through" capability is due to the fact that the original spectral prior knowledge base of the SIAM consists of a reference dictionary of spectral signatures in TOARF values, where relation $\mathrm{TOARF} \approx (\mathrm{SURF} + \mathrm{atmosphericnoise})$ holds, whereas traditional libraries of spectral signatures are in SURF values (measured at the ground level) exclusively, i.e., they are atmospheric noise-free. Well-known examples of reference dictionaries of spectral signatures in (atmospheric noise-free) SURF values, such as the U.S. Geological Survey (USGS) mineral and vegetation spectral libraries, the Johns Hopkins University spectral library and the Jet Propulsion Laboratory mineral spectral library [6]–[9], can be found in the existing literature, e.g., refer to [90] (p. 273) or in commercial software products [85]. Being provided with an (implicit) atmospheric noise model, the SIAM is expected to be robust to the presence of atmospheric effects. This means that SIAM does not consider preliminary atmospheric correction as mandatory because SIAM is knowledgeable on how to cope with RS data affected by atmospheric noise.

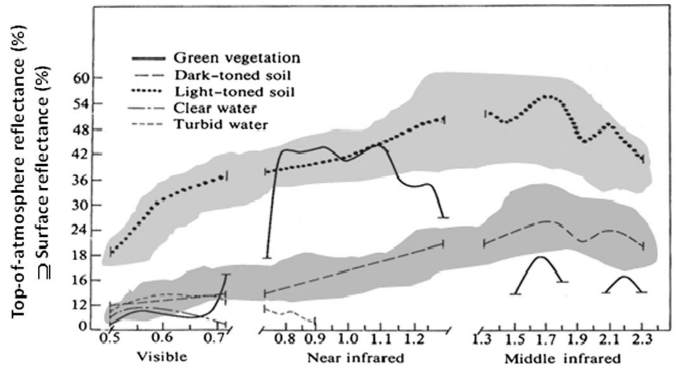

Fig. 3. Land cover (LC)-class specific families of spectral signatures in TOA reflectance (TOARF) values form buffer areas (envelopes) which include surface reflectance (SURF) values as a special case in clear sky and flat terrain conditions.

Unlike the SIAM reference dictionary of spectral signatures in TOARF values, the ATCOR-SPECL rule set has been developed starting from a prior knowledge base of reference spectral signatures in SURF values [6], [91], which means that the ATCOR-SPECL requires atmospheric correction as a mandatory preprocessing stage. In general, atmospheric correction is inherently poorly-conditioned and, therefore, difficult to solve. In practice, atmospheric correction requires user-supervision to become better posed for numerical treatment, also refer to Fig. 2 [6]–[9]. Although it requires SURF values as input data, the ATCOR-SPECL software product is expected to be able to cope with (to look-through) input images in TOARF values, when atmospheric effects are those typical of clear or very clear sky conditions and topographic effects are negligible, such that $\mathrm{TOARF} \approx \mathrm{SURF}$ [89].

2) First-Stage Output Semisymbolic Information Primitives: In a community-agreed ontology of the 4-D world-through-time (refer to Section II-C), e.g., in an LC or LCC map's legend (vocabulary), each ontological concept, e.g., each LC or LCC class name in the vocabulary, identifies a specific class of surface objects in the 4-D world-through-time featuring specific 4-D spatio-temporal properties, together with spectral (color) properties. In general, *LC class-specific spatio-temporal information dominates color information* [26] (refer to Section I), which is the reason why achromatic vision can be very successful despite the absence of color information.

In a preclassification map generated by the ATCOR-SPECL and SIAM software products from a single-date MS imagery, the map legend consists of a discrete and finite set of semisymbolic informational primitives, called color names, color-based inference categories, spectral-based semiconcepts, spectral categories or spectral endmembers, such as "vegetation," "bare soil or built-up," and "water or shadow" [10]–[19], [26]. Each spectral-based semiconcept can be mapped onto (matched with) one or more LC classes whose spectral properties can overlap, irrespective of spatio-temporal properties capable of disambiguating these LC classes (refer to Section I). In other words, spectral-based semiconcepts are single-date and pixel-specific, i.e., they ignore the (dominant) 4-D spatio-temporal information carried by LC classes, but exclusively investigate the (dominated) color properties of LC classes. As a consequence, the semantic meaning of a spectral-based semiconcept (e.g., "vegetation") is: 1) superior to zero, where zero is the semantic information conveyed by subsymbolic image features, i.e., image-objects (image-polygons) or, vice versa, image-contours (since image contour detection is the dual task of image segmentation and they are both poorly-posed [10]–[19]); and 2) equal or inferior to the semantic meaning of concepts in the attentive vision second stage, i.e., LC classes, e.g., "needle-leaf forest," belonging to a world model, namely, a spatio-temporal ontology of the 4-D world-through-time.

Hence, in general, one spectral-based semiconcept can be associated with none, one or many LC classes (refer to Section I). For example, spectral category "strong vegetation" can be linked to LC classes "grassland" or "agricultural field" or "forest," just like "*endmember fractions cannot always be inverted to unique class names*" ([26], p. 147). Analogously, one LC class can encompass different color discretization levels, e.g., the LC class "deciduous forest" can look like several tones of green equivalent to the SIAM's color quantization levels (spectral categories, color names) "strong vegetation," "average vegetation," and "dark vegetation." This means that, in general, a finite set of many-to-many associations holds between spectral-based semiconcepts in the (2-D) image domain and the reference LC classes belonging to a spatio-temporal ontology of the 4-D world-through-time [80]. Special cases of many-to-many inter-vocabulary relations are one-to-many, many-to-one and one-to-one relations. Many-to-many inter-legend relations convey mapping information because only all-to-all inter-legend "correct" entries do not (like if every spectral category were mapped onto all LC classes). For example, proposed in [80], an original Categorical Variable Pair Similarity Index (CVPSI) provides an estimated value, around 50%, of the degree of match between the SIAM's vocabulary and the LC class legend adopted by the USGS 2006 National Land Cover Data map, also refer to Fig. 1(c).

At a finer level of detail, SIAM delivers as output preclassification maps at various levels of color discretization, namely, fine, intermediate and coarse, where prior knowledge-based color quantization levels depend on the spectral resolution of the imaging sensor. At coarse granularity, SIAM's spectral categories belong to the following six parent spectral categories (also called super-categories) or major spectral endmembers: 1) "Clouds," 2) "Either snow or ice," 3) "Either water or shadow," 4) "Vegetation," equivalent to "either woody vegetation or cropland or grassland (herbaceous vegetation) or (shrub and brush) rangeland," 5) "Either bare soil or built-up," and 6) "Outliers."

These SIAM super-categories can be compared with the four reference endmembers, namely, "green vegetation," "nonphotosynthetic vegetation" (e.g., woody material on the ground together with dead or dying leaves), "soil," and "shadow," derived from laboratory surface reflectance spectra by Adams *et al.* in spectral mixture analysis [26].

Due to the presence of class "Outliers" ("Unknowns"), SIAM provides a mutually exclusive and totally exhaustive mapping of the input MS image into a discrete and finite vocabulary (legend) of color names, in line with the Congalton and Green requirements of a classification scheme [92]. It is noteworthy that, although the definition of a rejection rate is a well-known objective of any RS image classification system, e.g., refer to [26] and [90], RS image classifiers are often applied without any outlier detection strategy.

Similar considerations hold for the ATCOR-SPECL preclassifier, refer to the ATCOR-SPECL legend shown in Table I. For example, to identify information primitives of an ATCOR-SPECL's output map, the most recent ATCOR user guides, like [7] and [8], adopt the same term, "spectral categories," originally proposed in the SIAM literature to differentiate spectral-based semiconcepts from traditional LC classes [10]–[19]. According to [6]–[8], revised by Richter [91], the ATCOR-SPECL static decision-tree preclassifier consists of a sorted set of 19 spectral categories, including class "unknowns" (refer to Table I), in compliance with the Congalton and Green requirements of a classification scheme [92].

C. Algorithm Design

In [93], algorithm design is defined as "everything, but code." This definition is recalled to point out that, although they belong to the same family of spectral knowledge-based preclassifiers (refer to Section IV-A), capable of transforming subsymbolic observational data into semisymbolic spectral categories (refer to Section IV-B), the ATCOR-SPECL and SIAM software products are totally different in terms of decision-tree design, comprising both structural and procedural knowledge (refer to Section II-E), irrespective of implementation.

Sonka *et al.* describe aspects of image-object labeling through artificial intelligence in terms of syntactic pattern recognition ([30]; p. 285). In syntactic pattern recognition, the following considerations hold.

1) Elementary properties of the syntactically described objects from a given class are called primitives. Relations between objects may be modeled as hierarchical relational structures.
2) A class-specific description language is the set of all words that may be used to describe objects from one class, based on information primitives. For example, in written language, words of the language are constructed from letters and the set of all letters is called the alphabet. Letters are equivalent to information primitives and the words of the language are created from a collection of the alphabet's letters.

TABLE I

SPECTRAL RULES AND PSEUDO-COLORS OF THE LEGEND ADOPTED BY THE ATCOR-SPECL PRIOR KNOWLEDGE-BASED PRECLASSIFIER [6], [91]

Index	**Spectral categories**	**Spectral rule** (based on reflectance measured at Landsat TM central wave bands: b1 is located at 0.48 µm, b2 at 0.56 µm, b3 at 0.66 µm, b4 at 0.83 µm, b5 at 1.6 µm, and b7 at 2.2 µm)	**Pseudo-color**
1	Snow/ice	b4/b3 ≤ 1.3 AND b3 ≥ 0.2 AND b5 ≤ 0.12	
2	Cloud	b4 ≥ 0.25 AND 0.85 ≤ b1/b4 ≤ 1.15 AND b4/b5 ≥ 0.9 AND b5 ≥ 0.2	
3	Bright bare soil/ sand/cloud	b4 ≥ 0.15 AND 1.3 ≤ b4/b3 ≤ 3.0	
4	Dark bare soil	b4 ≥ 0.15 AND 1.3 ≤ b4/b3 ≤ 3.0 AND b2 ≤ 0.10	
5	Average vegetation	b4/b3 ≥ 3.0 AND (b2/b3 ≥ 0.8 OR b3 ≤ 0.15) AND 0.28 ≤ b4 ≤ 0.45	
6	Bright vegetation	b4/b3 ≥ 3.0 AND (b2/b3 ≥ 0.8 OR b3 ≤ 0.15) AND b4 ≥ 0.45	
7	Dark vegetation	b4/b3 ≥ 3.0 AND (b2/b3 ≥ 0.8 OR b3 ≤ 0.15) AND b3 ≤ 0.08 AND b4 ≤ 0.28	
8	Yellow vegetation	b4/b3 ≥ 2.0 AND b2 ≥_b3 AND b3 ≥ 8.0 AND b4/b5 ≥ 1.5[a]	
9	Mix of vegetation/ soil	2.0 ≤ b4/b3 ≤ 3.0 AND 0.05 ≤ b3 ≤ 0.15 AND b4 ≥ 0.15	
10	Asphalt/dark sand	b4/b3 ≤ 1.6 AND 0.05 ≤ b3 ≤ 0.20 AND 0.05 ≤ b4 ≤ 0.20[a] AND 0.05 ≤ b5 ≤ 0.25 AND b5/b4 ≥ 0.7[a]	
11	Sand/bare soil/ cloud	b4/b3 ≤ 2.0 AND b4 ≥ 0.15 AND b5 ≥ 0.15[a]	
12	Bright sand/bare soil/cloud	b4/b3 ≤ 2.0 AND b4 ≥ 0.15 AND (b4 ≥ 0.25[b] OR b5 ≥ 0.30[b])	
13	Dry vegetation/ soil	(1.7 ≤ b4/b3 ≤ 2.0 AND b4 ≥ 0.25[c]) OR (1.4 ≤ b4/b3 ≤ 2.0 AND b7/b5 ≤ 0.83[c])	
14	Sparse veg./soil	(1.4 ≤ b4/b3 ≤ 1.7 AND b4 ≥ 0.25[c]) OR (1.4 ≤ b4/b3 ≤ 2.0 AND b7/b5 ≤ 0.83 AND b5/b4 ≥ 1.2[c])	
15	Turbid water	b4 ≤ 0.11 AND b5 ≤ 0.05[a]	
16	Clear water	b4 ≤ 0.02 AND b5 ≤ 0.02[a]	
17	Clear water over sand	b3 ≥ 0.02 AND b3 ≥ b4 + 0.005 AND b5 ≤ 0.02[a]	
18	Shadow		
19	Not classified (outliers)		

[a] These expressions are optional and only used if b5 is present.
[b] Decision rule depends on presence of b5.
[c] Decision rule depends on presence of b7 [8].

3) A class-specific description grammar is the set of (substitution) rules that must be followed when words of the class-specific description language are constructed from letters. In other terms, each class consists only of objects whose syntactic description is syntactically correct according to the particular class grammar. In the written language example, although many words may be used together, only those which follow the correct grammar will form a coherent sentence.
4) Syntactic recognition is a process that looks for the class-specific grammar that can generate the syntactic word or phrase which describes an unknown object.
5) (Qualitative) syntactic object description should be used whenever (quantitative) statistical feature description is not able to represent the complexity of the target objects and/or when there are inter-object relations, like *part-of* or *subset-of*, difficult to learn from data by means of inductive data learning algorithms and that typically require significant human interaction to be identified.

In the aforementioned terminology of syntactic pattern recognition systems, both the ATCOR-SPECL and SIAM deductive decision-tree preclassifiers are built upon a physical knowledge base of families (envelops) of real-world spectral signatures per surface type (e.g., "bare soil or built-up"), so that a sorted set of land surface type-specific grammars (hierarchical decision-tree) is constructed.

In the SIAM software product, a spectral category-specific grammar is a combination of two information primitives capable of describing the family of spectral signatures belonging to that surface type (see [11] for full details). The first spectral primitive is the so-called "spectral rule" whose aim is to describe the shape of a buffer zone (envelope) of a surface type-specific family of spectral signatures in TOARF values, irrespective of intensity (see Fig. 2). In particular, a spectral rule defines a buffer zone of spectral tolerance, irrespective of the absolute intensity of spectral bands, by means of relational operators ($<$, $>$, $\leq$, $\geq$) between spectral bands. The second spectral primitive is a spectral fuzzy set (e.g., low, medium, and high) extracted from the intensity of scalar spectral variables, namely, spectral bands or spectral indexes. To recapitulate, a surface type-specific grammar is a combination of logical operators (AND, OR, NOT) with one or more spectral rules and/or one or more spectral fuzzy sets, capable of modeling the shape and the radiometric intensity of the surface type-specific MS envelope of spectral signatures [11].

Unlike SIAM, where a spectral category-specific grammar consists of a logical (AND, OR, NOT) combination of one or more spectral rules and spectral fuzzy sets [11], each ATCOR-SPECL's category-specific grammar consists of a single spectral rule per spectral category [6]–[8], see Table I.

Since the rule complexity of the SIAM expert system is superior to that of the ATCOR-SPECL, the former is expected to be more accurate than the latter at the cost of a higher implementation complexity and computation time.

To conclude this section, let us point out the algorithmic difference between the ATCOR-SPECL and SIAM prior knowledge-based preclassifiers and the popular spectral mixture analysis for MS image classification [26]. In spectral unmixing, the so-called (endmember) fraction categories are detected by category-specific boundaries established sequentially and in a particular order by an application developer in an E-dimensional measurement space, where E is the total number of reference endmembers, such that E is always less or equal than the number of spectral bands minus 1. For example, in the work of Adams *et al.* [26], dealing with 7-band Landsat images, the free number of spectral endmembers E is set equal to four, to allow the endmember space be rotated by the application developer on the computer screen to show any desired projection. On the contrary, the prior knowledge-based preclassification decision trees implemented in the ATCOR-SPECL and SIAM software products consist of dozens of prior knowledge-based category-specific grammars, whose inputs are spectral bands and spectral indexes, but never reference endmembers. Rather, the ATCOR-SPECL and SIAM expert systems, consisting of prior knowledge-based color discretization levels equivalent to data- and application-independent spectral endmembers, are suitable for automatic preclassification of hyperspectral images as a viable deductive alternative to state-of-the-art inductive algorithms for spectral endmember learning from hyperspectral data [104].

D. Implementation

The two ATCOR-SPECL and SIAM deductive decision-tree preclassifiers are totally different at the abstraction level of algorithm design (refer to Section IV-C), encompassing the list of category-specific grammars (structural knowledge [55]) and their order of presentation (procedural knowledge [55]). As a consequence, they are completely different at the implementation level of analysis.

According to [6]–[8], revised by Richter [91], the static decision-tree preclassifier currently implemented in the ATCOR-SPECL secondary software product consists of a sorted set of 19 spectral category-specific grammars (refer to Table I) which includes class "unknowns" (refer to Section IV-B2). In terms of semantic granularity the ATCOR-SPECL is coarser than the SIAM (vice versa, the semantic cardinality of the former is inferior to that of the latter), which means that the implementation complexity of the latter's decision tree is greater than that of the former (also refer to Section IV-C).

To the best of these authors' knowledge, *the SIAM software product is the first semisymbolic expert system* (refer to Section II-B)*, made available to the RS community for operational use in a RS-IUS preattentive vision first stage* (refer to Section II-C)*, capable of accomplishing multiscale image segmentation and multigranule image preclassification simultaneously, automatically and in near real-time* [10]–[19]. The extraction of a (subsymbolic) image segmentation map (where subsymbolic image-objects are identified as, say, segment 1, segment 2, etc.) from a binary or multilevel image (e.g., a thematic map) can be accomplished by a traditional well-posed (deterministic) automatic (requiring no user interaction) two-pass connected-component image labeling algorithm, e.g., refer to [30] (p. 197). In practice, a unique (subsymbolic) segmentation map can be generated from a multilevel image, like a thematic map, but the contrary does not hold, e.g., different thematic maps can generate the same segmentation map, i.e., no unequivocal thematic map can be inferred from a segmentation map [18], [19]. In other words, a realistic alternative to the (e.g., eCognition's) generation of an inherently poorly-conditioned, semiautomatic and site-specific multiscale segmentation map from an input subsymbolic MS image is the automatic well-posed generation of a multiscale segmentation map from a multilevel semisymbolic preclassification map, featuring several degrees of color discretization (e.g., fine, intermediate and coarse), which has been automatically generated by a prior knowledge-based multigranule preclassifier from an input MS image.

SIAM is implemented as an integrated system of six subsystems, including one "master" Landsat-like subsystem plus five "slave" (down-scale) subsystems, whose spectral resolution overlaps with Landsat's, but is inferior to Landsat's, refer to Table II. Noteworthy, the expression "Landsat-like MS image" adopted in this paper means: "an MS image whose spectral resolution mimics the spectral domain of the 7 bands of the Landsat family of imaging sensors," i.e., a spectral resolution where bands visible blue (B), visible green (G), visible red (R), near infra-red (NIR), medium infra-red 1 (MIR1), medium infra-red 2 (MIR2) and thermal infra-red (TIR) overlap (which does not mean coincide) with Landsat's.

The aforementioned SIAM's six subsystems are summarized in Table II. The output spectral categories detected at the fine, intermediate and coarse color discretization levels by the SIAM's six subsystems, described in Table II, are summarized in Table III.

With regard to the SIAM implementation, in [11] enough information is provided for the crisp L-SIAM implementation

TABLE II
LIST OF SPACEBORNE/AIRBORNE SENSORS ELIGIBLE FOR USE WITH THE SIAM SYSTEM OF SYSTEMS

SIAM system of systems		**B—(E)TM1, 0.45–0.52 (μm)**	**G—(E)TM2, 0.52–0.60 (μm)**	**R—(E)TM3, 0.63–0.69 (μm)**	**NIR—(E)TM4, 0.76–0.90 (μm)**	**MIR1—(E)TM5, 1.55–1.75 (μm)**	**MIR2—(E)TM7, 2.08–2.35 (μm)**	**TIR—(E)TM6, 10.4–12.5 (μm)**	**SR (m)**	**Rad. Cal. Y/N, C/I**	**Pan SR (m)**	**Notes**
L-SIAM, Input bands: 7 – B, G, R, NIR, MIR1, MIR2, and TIR. Output Sp. Cat.: 96/48/18	**Landsat-4/-5 TM**	×	×	×	×	×	×	×	30	Y-C		Refer to Table I in [11]
	Landsat-7 ETM+	×	×	×	×	×	×	×	30	Y-C	15	Same as above.
	Landsat-8 OLI+TIRS	×	×	×	×	×	×	×	30	Y-C	15	
	MODIS	×	×	×	×	×	×	×	250, 500, 1000	Y-C		Same as above.
	ASTER		×	×	×	×	×	×	15-30	Y-C		Same as above.
	CBERS-2B	×	×	×	×	×	×	×		N		
	APEX	×	×	×	×	×	×		1.8	Y		Airborne hyperspectral, 285 bands
	AVIRIS	×	×	×	×	×	×		e.g., 20	Y-?		Airborne hyperspectral, 224 bands, managed by Jet Propulsion Laboratory (JPL)
	MIVIS	×	×	×	×	×	×	×	e.g., 1.64	Y-?		Airborne hyperspectral, 102 bands, managed by CNR, Italy
	Sentinel-2 MSI	×	×	×	×	×	×		10 (B, G, R, NIR), 20 (MIR1, MIR2)	?		13 bands, from VIS to MIR. To be launched in 2015?
	Sentinel-3 SLSTR		×	×	×	×	×	×	500	?		9 bands, from VIS to TIR + 2 (active fire). To be launched in 2015?
	WorldView-3	×	×	×	×	×	×		MS: 1.24, SWIR: 3.7	Y-C	0.3	16 bands, from VIS to SWIR. Launched in Aug. 2014.
S-SIAM, Input bands: 4—G, R, NIR, MIR1. Output Sp. Cat.: 68/40/15	**SPOT-4 HRVIR**		×	×	×	×			20	Y-I	10	Refer to Table II in [11].
	SPOT-5 HRG		×	×	×	×			10	Y-I	2.5–5	Same as above.
	SPOT-4/-5 VMI		×	×	×	×			1100	Y-I		Same as above.
	IRS-1C/-1D LISS-III		×	×	×	×			23.5	Y-I		
	IRS-P6 LISS-III		×	×	×	×			23.5	Y-I		
	IRS-P6 AWiFS		×	×	×	×			56	Y-I		

TABLE II
CONTINUED

AV-SIAM, Input bands: 4 —R, NIR, MIR1, TIR. Output Sp. Cat.: 83/43/17	**NOAA AVHRR**			×	×	×		×	1100	Y		Refer to Table II in [11].
	MSG			×	×	×		×	3000	Y		Same as above.
	NASA-NOAA NPP VIIRS			×	×	×	×	×	375	Y-C		
AA-SIAM, Input bands: 5 —G, R, NIR, MIR1, TIR. Output Sp. Cat.: 83/43/17	**ENVISAT AATSR**		×	×	×	×		×	1000	Y		Same as above.
	ERS-2 ATSR-2		×	×	×	×		×	1000	Y		
Q-SIAM, Input bands: 4 —B, G, R, NIR. Output Sp. Cat.: 61/28/12	**IKONOS-2**	×	×	×	×				4	Y-C	1	
	QuickBird-2	×	×	×	×				2.4	Y-C	0.61	
	GeoEye-1	×	×	×	×				1.64	Y	0.41	
	OrbView-3	×	×	×	×				4	N	1	
	SPOT-6/7	×	×	×	×				6	Y-I	1.5	
	Pleaides-1A/1B	×	×	×	×				2	Y-I	0.5	
	RapidEye-1 to -5	×	×	×	×				6.5	Y-I		
	ALOS AVNIR-2	×	×	×	×				10	Y-C		
	KOMPSAT-2	×	×	×	×				4	N	1	
	TopSat	×	×	×	×				5	N	2.5	
	FORMOSAT-2	×	×	×	×				8	Y-?	2	
	Huan Jing satellite constellation, HJ-1A / HJ-1B, payload: WVC.	×	×	×	×				30	Y-C		Wide View CCD cameras (WVC).
	ENVISAT MERIS	×	×	×	×				300	Y-?		Super-spectral, 15 bands
	Sentinel-3 OLCI	×	×	×	×				300, 1200			Super-spectral, 21 bands. To be launched in 2015?
	Leica ADS-40/80	×	×	×	×				0.25	Y-?	0.25	Airborne, 4 bands + PAN
D-SIAM, Input bands: 3 —G, R, NIR. Output Sp. Cat.: 61/28/12	**Landsat-1/-2/-3/-4/-5 MSS**		×	×	×				79	Y-C		
	IRS-P6 LISS-IV		×	×	×				5.8	Y-I		
	SPOT-1/-2/-3 HRV		×	×	×				20	Y-I	10	
	DMC		×	×	×				22-32	Y-C		

Acronyms: Y, Yes; N, No; C, Complete; I, Incomplete (radiometric calibration offset parameters are set to zero); (E)TM, (Enhanced) Thematic Mapperl; B, Blue; G, Green; R, Red; NIR, Near Infra-Red; MIR, Medium Infra-Red; TIR, Thermal Infra-Red; SR, Spatial Resolution; and Pan, Panchromatic.
Adopted acronyms: SPOT, Satellite Pour l'Observation de la Terre; NOAA, National Oceanic and Atmospheric Administration (NOAA); AVHRR, Advanced Very High Resolution Radiometer; AATSR, ENVISAT Advanced Along-Track Scanning Radiometer; Q, QuickBird; DMC, Disaster Monitoring Constellation.
Column highlight color: Blue columns are related to visible channels typical of water and haze; Green column identify the NIR band, typical of vegetation; Brown columns are related to MIR channels, characteristic of bare soils; and Red column: TIR channel, useful to detect fire.

to be reproduced. The down-scale S-SIAM, AV-SIAM and Q-SIAM versions, generated from the "master" L-SIAM implementation (refer to Table II), are described in [12]–[14]. In [17], the crisp-to-fuzzy SIAM transformation is explained in detail. It is noteworthy that since its first 2006 release presented in [11], L-SIAM has increased its number of output spectral categories from 46 to 96 (see Table II). This progressive, but slow, increase in the number of spectral categories detected by the sequence of "master" L-SIAM implementations proposed to the RS literature in recent years shows that, in line with theory [45], [55] (refer to Section II-B), there is a slow "learning curve" in the development and fine-tuning of physical models, such as the ATCOR-SPECL and SIAM prior knowledge-based preclassifiers.

TABLE III
SIAM System of Six Subsystems

SIAM	Input bands (B: Blue, G: Green, R: Red, NIR: Near Infra-Red, MIR: Medium IR, TIR: Thermal IR)	Preliminary classification map output products: number of output spectral categories.			
		Fine semantic granularity	Intermediate semantic granularity	Coarse semantic granularity	Inter-sensor semantic granularity (*)
L-SIAM	7—B, G, R, NIR, MIR1, MIR2, TIR	96	48	18	33
S-SIAM	4—G, R, NIR, MIR1	68	40	15	
AV-SIAM	4—R, NIR, MIR1, TIR	83	43	17	
AA-SIAM	5—G, R, NIR, MIR1, TIR	83	43	17	
Q-SIAM	4—B, G, R, NIR	61	28	12	
D-SIAM	3—G, R, NIR	61	28	12	

*Employed in sensor-independent bitemporal LCC detection.
Summary of input bands and output spectral categories reported in Table II.

V. Conclusion

In compliance with the QA4EO guidelines, the goal of this paper is to provide a theoretical comparison and an experimental quality assessment of two operational (ready-for-use) expert systems (prior knowledge-based nonadaptive decision trees) for automatic near real-time preattentional classification and segmentation of spaceborne/airborne MS images: the SIAM software product and the SPECL secondary product of the ATCOR commercial software toolbox. Rather than as standalone systems, these two alternative prior knowledge-based preclassifiers in operating mode are eligible for use in the preattentive vision first stage of a novel hybrid (combined deductive and inductive) RS-IUS architecture, proposed to the RS community in recent years [10]–[20].

For the sake of simplicity, this paper is split into two: Part 1—Theory, proposed herein, and Part 2—Experimental results, already published elsewhere [20].

The original contribution of the present Part 1 is threefold. First, it provides Part 2 with an interdisciplinary terminology and a theoretical background encompassing multiple disciplines, like philosophical hermeneutics, machine learning, artificial intelligence, computer vision, human vision and RS. Second, it highlights the relevant degrees of novelty of the ATCOR-SPECL and SIAM prior knowledge-based preclassifiers at the four levels of understanding of an information processing system, namely, system design, knowledge/information representation, algorithms and implementation. Third, it requires that a minimum set of community-agreed complementary independent metrological/statistically-based QIOs must be estimated from a RS-IUS in operating mode, to comply with the principles of statistics, the QA4EO guidelines [2] and the Committee on EO Satellites (CEOS) land product accuracy validation criteria [3]. In particular, sample QIs of the ATCOR-SPECL and SIAM prior knowledge-based preclassifiers, to be collected in Part 2 of this paper, must be: 1) statistically significant, i.e., provided with a degree of uncertainty in measurement, and 2) statistically valid (consistent), i.e., representative of the entire population being sampled, which requires the implementation of a probability sampling protocol [82], [83]. Noteworthy, these basic sample statistic requirements should not be considered either trivial or obvious. For example, they are almost never satisfied in the RS common practice. As a consequence, to date, QIOs of existing RS-IUSs, including mapping accuracy, in addition to degree of automation, efficiency, robustness, scalability, timeliness and costs, remain largely unknown in statistical terms.

The conclusion of the present Part 1 of this paper is that the proposed comparison of the ATCOR-SPECL and SIAM software products in operating mode, accomplished in Part 2, can be considered appropriate, well-timed and of potential interest to a wide portion of the RS community.

Acknowledgment

A. Baraldi would like to thank R. Capurro for his hospitality, patience, politeness, and open-mindedness. The authors would like to thank L. Boschetti for his support, who was at the University of Maryland, College Park, MD, USA, and is now with the College of Natural Resources, University of Idaho, Moscow, ID, USA, and the Editor-in-Chief, Associate Editor, and reviewers for their competence, patience, and willingness to help.

References

[1] Group on Earth Observations (GEO). (2005, Feb. 16). *The Global Earth Observation System of Systems (GEOSS) 10-Year Implementation Plan* [Online]. Available: http://www.earthobservations.org/docs/10-Year%20Implementation%20Plan.pdf, accessed on Nov. 15, 2012.

[2] Group on Earth Observations (GEO)–Committee on Earth Observation Satellites (CEOS). (2010). *A Quality Assurance Framework for Earth Observation, Version 4.0* [Online]. Available: http://qa4eo.org/docs/QA4EO_Principles_v4.0.pdf, accessed on Nov. 15, 2012.

[3] Group on Earth Observations (GEO)-Committee on Earth Observation Satellites (CEOS) - CEOS Working Group on Calibration and Validation - Land Product Validation Subgroup. (2014, Aug. 19). *Land Product Validation Mission* [Online]. Available: http://lpvs.gsfc.nasa.gov/, accessed on Sep. 3, 2014.

[4] G. Gutman *et al.*, Eds., *Land Change Science*. Norwell, MA, USA: Kluwer, 2004.

[5] D. Marr, *Vision*. San Francisco, CA, USA: W.H. Freeman, 1982.

[6] W. Dorigo, R. Richter, F. Baret, R. Bamler, and W. Wagner, "Enhanced automated canopy characterization from hyperspectral data by a novel two step radiative transfer model inversion approach," *Remote Sens.*, vol. 1, pp. 1139–1170, 2009, doi: 10.3390/rs1041139.

[7] R. Richter and D. Schläpfer, *Atmospheric/topographic correction for satellite imagery, ATCOR-2/3 User Guide, Version 8.2.1*, DLR-IB 565-01/13. Wessling, Germany: DLR-German Aerospace Center, Feb. 2013 [Online]. Available: http://www.rese.ch/pdf/atcor3_manual.pdf, accessed on Sep. 3, 2014.

[8] R. Richter and D. Schläpfer, *Atmospheric/topographic correction for airborne imagery, ATCOR-4 User Guide, Version 6.2.1*, DLR-IB 565-02/13. Wessling, Germany, DLR-German Aerospace Center, Feb. 2013 [Online]. Available: http://www.rese.ch/pdf/atcor4_manual.pdf, accessed on Sep. 3, 2014.

[9] R. Richter, D. Schläpfer, and A. Hueni, "Recent developments in operational atmospheric and radiometric correction of hyperspectral imagery," in *Proc. 6th EARSeL SIG IS Workshop*, Tel Aviv, Israel, Mar. 16–19, 2009, pp. 1–6 [Online]. Available: http://www.earsel6th.tau.ac.il/∼earsel6/CD/PDF/earsel-PROCEEDINGS/3054%20Schl%20pfer.pdf, accessed on Sep. 3, 2014.

[10] 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 Observ. Remote Sens.*, vol. 2, no. 2, pp. 104–134, Jun. 2009.

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

[12] A. Baraldi *et al.*, "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.

[13] A. Baraldi *et al.*, "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.

[14] A. Baraldi *et al.*, "Corrections to '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," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 3, p. 1635, Mar. 2010.

[15] 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.

[16] A. Baraldi, T. Wassenaar, and S. Kay, "Operational performance of an automatic preliminary spectral rule-based decision-tree classifier of spaceborne very high resolution optical images," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 9, pp. 3482–3502, Sep. 2010.

[17] A. Baraldi, "Fuzzification of a crisp near-real-time operational automatic spectral-rule-based decision-tree preliminary classifier of multisource multispectral remotely sensed images," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 6, pp. 2113–2134, Jun. 2011.

[18] A. Baraldi and L. Boschetti, "Operational automatic remote sensing image understanding systems: Beyond geographic object-based and object-oriented image analysis (GEOBIA/GEOOIA)—Part 1: Introduction," *Remote Sens.*, vol. 4, no. 9, pp. 2694–2735, 2012, doi: 10.3390/rs4092694.

[19] A. Baraldi and L. Boschetti, "Operational automatic remote sensing image understanding systems: Beyond geographic object-based and object-oriented image analysis (GEOBIA/GEOOIA)—Part 2: Novel system architecture, information/knowledge representation, algorithm design and implementation," *Remote Sens.*, vol. 4, no. 9, pp. 2768–2817, 2012, doi: 10.3390/rs4092768.

[20] A. Baraldi, M. Humber, and L. Boschetti, "Quality assessment of preclassification maps automatically generated from spaceborne/airborne multi-spectral images by the Satellite Image Automatic Mapper and Atmospheric/Topographic Correction-Spectral Classification software products: Part 2—Experimental results," *Remote Sens.*, vol. 5, pp. 5209–5264, Oct. 2013.

[21] R. Capurro and B. Hjørland, "The concept of information," *Ann. Rev. Info. Sci. Tech.*, vol. 37, no. 1, pp. 343–411, 2003 [Online]. Available: http://www.capurro.de/infoconcept.html, accessed on 3 Sep. 2014.

[22] R. Capurro, "Hermeneutics and the phenomenon of information," in *Metaphysics, Epistemology, and Technology. Research in Philosophy and Technology*. Amsterdam, The Netherlands: JAI/Elsevier, 2000, vol. 19, pp. 79–85.

[23] E. Diamant, "Machine learning: When and where the horses went astray?" in *Machine Learning*, Y. Zhang, Ed. Rijeka, Croatia: InTech Open Access, 2010, pp. 1–18.

[24] V. Cherkassky and F. Mulier, *Learning From Data: Concepts, Theory, and Methods*. Hoboken, NJ, USA: Wiley, 1998.

[25] C. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pp. 379–423, 623–656, 1948.

[26] J. B. Adams *et al.*, "Classification of multispectral images based on fractions of endmembers: Application to land-cover change in the Brazilian Amazon," *Remote Sens. Environ.*, vol. 52, pp. 137–154, 1995.

[27] Yale Center for Earth Observation. (2009, Oct. 11). *South West Asia Project - Landcover Classification*. Yale University, Yale, NH, USA [Online]. Available: http://www.yale.edu/ceo/Projects/swap/landcover/Unsupervised_classification.htm, accessed on Sep. 3, 2014.

[28] J. Cihlar, R. Latifovic, and J. Beaubien, "A comparison of clustering strategies for unsupervised classification," *Can. J. Remote Sens.*, vol. 26, pp. 446–454, 2000.

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

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

[31] P. Mather, *Computer Processing of Remotely-Sensed Images—An Introduction*. Hoboken, NJ, USA: Wiley, 1994.

[32] 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.

[33] M. Baatz and A. Schäpe, "Multiresolution segmentation: An optimization approach for high quality multi-scale image segmentation," *J. Photogramm. Remote Sens.*, vol. 58, no. 3, pp. 12–23, 2000.

[34] M. Baatz, C. Hoffmann, and G. Willhauck, "Progressing from object-based to object-oriented image analysis," in *Object-Based Image Analysis—Spatial Concepts for Knowledge-Driven Remote Sensing Applications*, T. Blaschke, S. Lang and G. J. Hay, Eds. New York, NY, USA: Springer, 2008, ch. 1.4, pp. 29–42.

[35] S. Tatiraju and A. Mehta, "Image Segmentation using k-means clustering, EM and normalized cuts," Univ. California, Irvine, CA, USA, Tech. Rep., 2008, pp. 1–7 [Online]. Available: http://www.ics.uci.edu/∼dramanan/teaching/winter08/projects/avim˙report.pdf, accessed on Sep. 3, 2014.

[36] N. R. Pal and S. K. Pal, "A review on image segmentation techniques," *Pattern Recognit.*, vol. 26, no. 9, pp. 1277–1294, 1993.

[37] A. Baraldi and E. Alpaydin, "Constructive feedforward ART clustering networks—Part I," *IEEE Trans. Neural Netw.*, vol. 13, no. 3, pp. 645–661, May 2002.

[38] A. Baraldi and E. Alpaydin, "Constructive feedforward ART clustering networks—Part II," *IEEE Trans. Neural Netw.*, vol. 13, no. 3, pp. 662–677, May 2002.

[39] B. Fritzke, "Some competitive learning methods," Bochum, Germany: Ruhr-Universität Bochum, 1997 [Online]. Available: http://www.ki.inf.tu-dresden.de/∼fritzke/JavaPaper

[40] E. Backer and A. K. Jain, "A clustering performance measure based on fuzzy set decomposition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-3, no. 1, pp. 66–75, Jan. 1981.

[41] C. Persello and L. Bruzzone, "A novel protocol for accuracy assessment in classification of very high resolution images," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 3, pp. 1232–1244, Mar. 2010.

[42] J. Hadamard, "Sur les problemes aux derivees partielles et leur signification physique," *Princeton Univ. Bull.*, vol. 13, pp. 49–52, 1902.

[43] N. Hunt and S. Tyrrell. (2004, Jun. 10). *Stratified Sampling*. Coventry University, U.K. [Online]. Available: http://nestor.coventry.ac.uk/∼nhunt/meths/strati.html, accessed on 3 Sep. 2014.

[44] P. Zamperoni, "Plus ça va, moins ça va," *Pattern Recognit. Lett.*, vol. 17, pp. 671–677, 1996.

[45] S. Liang, *Quantitative Remote Sensing of Land Surfaces*. Hoboken, NJ, USA: Wiley, 2004.

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

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

[48] T. F. Cootes and C. J. Taylor, *Statistical Models of Appearance for Computer Vision, Imaging Science and Biomedical Engineering*. Manchester, U.K.: Univ. Manchester, 2004.

[49] D. K. Hall, G. A. Riggs, V. V. Salomonson, N. E. Di Girolamo, and K. J. Bayr, "MODIS snow-cover products," *Remote Sens. Environ.*, vol. 83, pp. 181–194, 2002.

[50] Y. Luo, A. P. Trishchenko, and K. V. Khlopenkov, "Developing clear-sky, cloud and cloud shadow mask for producing clear-sky composites at 250-meter spatial resolution for the seven MODIS land bands over Canada and North America," *Remote Sens. Environ.*, vol. 112, pp. 4167–4185, 2008.

[51] A. K. Jain, "Knowledge-based segmentation of Landsat images," *IEEE Trans. Geosci. Remote Sens.*, vol. 29, no. 2, pp. 222–232, Mar. 1991.

[52] M. Avcia and Z. Akyurek, "A hierarchical classification of Landsat TM imagery for landcover mapping," in *Proc. Int. Soc. Photogramm. Remote Sens. ISPRS XXXV Congr. WG IV/6*, Istanbul, Turkey, Jul. 12–23, 2004, pp. 1–6 [Online]. Available: http://www.isprs.org/proceedings/XXXV/congress/comm4/papers/402.pdf, accessed on Sep. 3, 2014.

[53] L. Moller-Jensen, "Classification of urban land cover based on expert systems, object models and texture," *Environ. Urban Syst.*, vol. 21, no. 3/4, pp. 291–302, 1997.

[54] A. K. Skidmore, "An expert system classifies Eucalypt forest types using TM data and a digital terrain model," *Photogramm. Eng. Remote Sens.*, vol. 55, no. 10, pp 1449–1464, 1989.

[55] S. Lang, "Object-based image analysis for remote sensing applications: Modeling reality—Dealing with complexity," in *Object-Based Image Analysis–Spatial Concepts for Knowledge-Driven Remote Sensing Applications*, T. Blaschke, S. Lang and G. J. Hay, Eds. New York, NY, USA: Springer, 2008, ch. 1.1, pp. 3–27.

[56] G. J. Hay and G. Castilla, "Object-based image analysis: Strengths, weaknesses, opportunities and threats (SWOT)," in *Proc. 1st Int. Conf. Object-Based Image Anal. (OBIA)*, Salzburg, Austria, Jul. 2006, pp. 1–3.

[57] G. J. Hay and G. Castilla, "Geographic object-based image analysis (GEOBIA): A new name for a new discipline," in *Object-Based Image Analysis: Spatial Concepts for Knowledge-Driven Remote Sensing Applications*, T. Blaschke, S. Lang and G. J. Hay, Eds. New York, NY, USA: Springer, 2008, ch. 1.4, pp. 81–92.

[58] Definiens Imaging GmbH, *eCognition Elements User Guide 4*. Munich, Germany: Definiens Imaging GmbH, 2004.

[59] Definiens AG, *Developer 8 Reference Book*. Munich, Germany: Definiens AG, 2011.

[60] Trimble. (2014). *eCognition Developer 9.0* [Online]. Available: http://www.ecognition.com/products/ecognition-developer, accessed on Sep. 3, 2014.

[61] 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. 2013.

[62] 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.

[63] R. R. Irish, J. L. Barker, S. N. Goward, and T. Arvidson, "Characterization of the Landsat-7 ETM+ automated cloud-cover assessment (ACCA) algorithm," *Photogramm. Eng. Remote Sens.*, vol. 72, no. 10, pp. 1179–1188, Oct. 2006.

[64] Q. Iqbal and J. K. Aggarwal, "Image retrieval via isotropic and anisotropic mappings," in *Proc. IAPR Workshop Pattern Recognit. Inf. Syst.*, Setubal, Portugal, Jul. 2001, pp. 34–49.

[65] M. F. Goodchild, M. Yuan, and T. J. Cova, "Towards a general theory of geographic representation in GIS," *Int. J. Geogr. Inf. Sci.*, vol. 21, no. 3, pp. 239–260, 2007.

[66] P. Lüscher, D. Burghardt, and R. Weibel, "Ontology-driven enrichment of spatial databases," in *Proc. 10th ICA Workshop Gen. Multiple Represent.*, Moscow, Russia, Aug. 2007, pp. 1–13 [Online]. Available: http://www.geo.uzh.ch/∼luescher/publications/luescher_genws2007.pdf, accessed on Sep. 3, 2014.

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

[68] P. Gouras, "Color vision," in *Principles of Neural Science*, E. Kandel and J. Schwartz, Eds. Norwalk, CT, USA: Appleton and Lange, 1991, pp. 467–479.

[69] E. R. Kandel, "Perception of motion, depth and form," in *Principles of Neural Science*, E. Kandel and J. Schwartz, Eds. Norwalk, CT, USA: Appleton and Lange, 1991, pp. 441–466.

[70] M. Carandini and D. Ferster, "Membrane potential and firing rate in cat primary visual cortex," *J. Neuroscience*, vol. 20, pp. 470–484, 2000.

[71] J. Mutch and D. Lowe, "Object class recognition and localization using sparse features with limited receptive fields," *Int. J. Comput. Vis.*, vol. 80, pp. 45–57, 2008.

[72] T. N. Wiesel and D. H. Hubel, "Spatial and chromatic interactions in the lateral geniculate body of the rhesus monkey," *J. Neurophys.*, vol. 29, pp. 1115–1156, 1966.

[73] A. Jain and G. Healey, "A multiscale representation including opponent color features for texture recognition," *IEEE Trans. Image Proc.*, vol. 7, no. 1, pp. 124–128, Jan. 1998.

[74] T. Martinetz and K. Schulten, "Topology representing networks," *Neural Networks*, vol. 7, no. 3, pp. 507–522, 1994.

[75] S. P. Vecera and M. J. Farah, "Is visual image segmentation a bottom-up or an interactive process?" *Percept. Psychophys.*, vol. 59, 1280–1296, 1997.

[76] J. D. McCafferty, *Human and Machine Vision, Computing Perceptual Organization*. Chichester, U.K.: Ellis Horwood Limited, 1990.

[77] J. Yang and R. S. Wang, "Classified road detection from satellite images based on perceptual organization," *Int. J. Remote Sens.*, vol. 28, pp. 4651–4669, 2007.

[78] Wikipedia. (2014, Sep. 1). *Big data* [Online]. Available: http://en.wikipedia.org/wiki/Big_data, accessed on Sep. 3, 2014.

[79] L. Alparone, S. Baronti, A. Garzelli, and F. Nencini, "A global quality measurement of pan-sharpened multispectral imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 1, no. 4, pp. 313–317, Oct. 2004.

[80] A. Baraldi, L. Boschetti, and M. Humber, "Probability sampling protocol for thematic and spatial quality assessments of classification maps generated from spaceborne/airborne very high resolution images," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 1, Part: 2, pp. 701–760, Jan. 2014.

[81] Y. Y. Boykov and M. P. Jolly, "Interactive graph cuts for optimal boundary and region segmentation of objects in N-D images," in *Proc. Int. Conf. Comput. Vis.*, Vancouver, BC, Canada, Jul. 2001, vol. I, pp. 105–112.

[82] S. V. Stehman and R. L. Czaplewski, "Design and analysis for thematic map accuracy assessment: Fundamental principles," *Remote Sens. Environ.*, vol. 64, pp. 331–344, 1998.

[83] W. S. Overton and S. V. Stehman, "The Horvitz–Thompson theorem as a unifying perspective for probability sampling: With examples from natural resource sampling," *Amer. Statistician*, vol. 49, no. 3, pp. 261–268, Aug. 1995.

[84] M. A. Wulder, J. G. Masek, W. B. Cohen, T. R. Loveland, and C. E. Woodcock, "Opening the archive: How free data has enabled the science and monitoring promise of Landsat," *Remote Sens. Environ.*, vol. 122, pp. 2–10, July 2012.

[85] ENVI 4.3, User Manual, ITT Industries, Inc., 2006.

[86] E. Vermote, D. Tanré, J. L. Deuzé, M. Herman, and J. J. Morcrette. (1997, Jul.). *Second Simulation of the Satellite Signal in the Solar Spectrum (6S), 6S User Guide Version 2* [Online]. Available: http://www.rsgis.ait.ac.th/∼honda/textbooks/advrs/6smanv2.0_P1.pdf, accessed on Jan. 10, 2012.

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

[88] 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.

[89] P. S. Chavez, "An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data," *Remote Sens. Environ.*, vol. 24, pp. 459–479, 1988.

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

[91] R. Richter, DLR-German Aerospace Center, personal communication to A. Baraldi.

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

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

[94] Z. Wang and A. C. Bovik, "A universal image quality index," *IEEE Signal Process. Lett.*, vol. 9, no. 3, pp. 81–84, Mar. 2002.

[95] G. Schaepman-Strub, M. E. Schaepman, T. H. Painter, S. Dangel, and J. V. Martonchik, "Reflectance quantities in optical remote sensing-definitions and case studies," *Remote Sens. Environ.*, vol. 103, pp. 27–42, 2006.

[96] European Commission Joint Research Center (JRC)–Institute for Environment and Sustainability (IES). *IMAGE2000 and CORINE Land Cover (CLC) 2000*. Ispra, Italy: JRC–IES, 2005, EUR 21757, ISBN 92-894-9862-5.

[97] European Commission Joint Research Center, *The IMAGE-2006 Mosaic Project*, EC-JRC: Ispra, Italy, EUR 23755, 2008 [Online]. Available: http://publications.jrc.ec.europa.eu/repository/bitstream/111111111/22218/1/lbna23755enn.pdf, accessed on Sep. 3, 2014.

[98] 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 N. Bahram, Eds. Boca Raton, FL, USA: CRC Press, 1992, pp. 309–327.

[99] H. R. Wilson and J. R. Bergen, "A four mechanism model for threshold spatial vision," *Vis. Res.*, vol. 19, no. 1, pp. 19–32, 1979.

[100] H. du Buf and J. Rodrigues, "Image morphology: From perception to rendering," in *IMAGE—Computational Visualistics and Picture Morphology*, 2007 [Online]. Available: http://w3.ualg.pt/∼jrodrig/papers_pdf/IMAGE07.pdf, accessed on Aug. 16, 2014.

[101] D. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, 2004.

[102] J. van de Weijer, C. Schmid, J. Verbeek, and D. Larlus, "Learning color names for real-world applications," *IEEE Trans. Image Process.*, vol. 18, no. 7, pp. 1512–1523, Jul. 2009.

[103] R. Benavente, M. Vanrell, and P. Baldrich, "Parametric fuzzy sets for automatic color naming," *J. Opt. Soc. Amer. A*, vol. 25, pp. 2582–2593, 2008.

[104] A. Plaza, J. Plaza, and G. Martin, "Spatial-spectral endmember extraction from hyperspectral imagery using multi-band morphology and volume optimization," in *Proc. 16th IEEE Int. Conf. Image Process. (ICIP)*, Cairo, Egypt, Nov. 2009, pp. 3721–3724.

Andrea Baraldi received the Laurea (M.S.) degree in electronic engineering from the University of Bologna, Bologna, Italy, in 1989, and a Master degree in software engineering from the University of Padova, Padova, Italy, and Purdue University, West Lafayette, IN, USA, in 1994.

From 1989 to 1990, he worked as a Research Associate at the Centro di Studio per l'Interazione Operatore–Calcolatore, National Research Council (CNR), Bologna, Italy, and served in the army with the Istituto Geografico Militare, Florence, Italy. He worked as a Consultant with the European Space Agency–European Space Research Institute, Frascati, Italy, from 1991 to 1993. From 1997 to 1999 he was a PostDoctoral Fellow in Artificial Intelligence with the International Computer Science Institute, Berkeley, CA, USA. He worked as a PostDoctoral Researcher with the Global Vegetation Monitoring Unit, the Institute for Environment and Sustainability (IES), the European Commission Joint Research Centre (JRC), Ispra, Italy, from 2000 to 2002. From 2005 to 2009, he was with the IES-Spatial Data Infrastructure, JRC. In 2009, he founded Baraldi Consultancy in Remote Sensing, a one-man company located in Modena, Italy. Since his master thesis until 2010 he continued his collaboration with the Istituto di Scienze dell'Atmosfera e del Clima, CNR, Bologna, and the Istituto di Studi di Sistemi Intelligenti per l'Automazione, CNR, Bari, Italy. From 2010 to 2013, he joined the Department of Geographical Studies, University of Maryland, College Park, MD, USA, where, to date, he is an Adjunct Associate Professor. In 2014, he was appointed with a Senior Scientist Fellowship at the German Aerospace Center (DLR), Oberpfaffenhofen, Germany. From June 2014 to October 2104, he was a Visiting Scientist at the Department of Geoinformatics—Z_GIS, University of Salzburg, Salzburg, Austria. His research interests include image preprocessing and understanding, with special emphasis on the development of operational automatic hierarchical multisource multiresolution spaceborne/airborne image understanding systems, consistent with human vision.

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

Michael L. Humber received the M.P.S. degree in geospatial information sciences in 2013, and is pursuing the Ph.D. degree at the Department of Geographical Sciences, University of Maryland, College Park, MD, USA.

He is a Research Assistant with the University of Maryland. Since 2010, he has been employed with the Department of Geographical Sciences Research Laboratory. His research interests include processing and validation of the MODIS Burned Area Products, validation of maps generated from moderate to very high resolution imagery, global agricultural monitoring, and web mapping in the context of facilitated volunteered geographic information (VGI) applications.