GEOINFORMATICS | RESEARCH ARTICLE

GEO-CEOS stage 4 validation of the Satellite Image Automatic Mapper lightweight computer program for ESA Earth observation level 2 product generation - Part 1: Theory

Andrea Baraldi, Michael Laurence Humber, Dirk Tiede and Stefan Lang

Cogent Geoscience (2018), 4: 1467357
GEO-INFORMATICS | RESEARCH ARTICLE

GEO-CEOS stage 4 validation of the Satellite Image Automatic Mapper lightweight computer program for ESA Earth observation level 2 product generation - Part 1: Theory

Andrea Baraldi1,2,3,4*, Michael Laurence Humber2, Dirk Tiede3 and Stefan Lang3

Abstract: ESA defines as Earth Observation (EO) Level 2 information product a single-date multi-spectral (MS) image corrected for atmospheric, adjacency and topographic effects, stacked with its data-derived scene classification map (SCM), whose legend includes quality layers cloud and cloud-shadow. No ESA EO Level 2 product has ever been systematically generated at the ground segment. To fill the information gap from EO big data to ESA EO Level 2 product in compliance with the GEO-CEOS stage 4 validation (Val) guidelines, an off-the-shelf Satellite Image Automatic Mapper (SIAM) lightweight computer program was validated by independent means on an annual 30 m resolution Web-Enabled Landsat Data (WELD) image composite time-series of the conterminous U.S. (CONUS) for the years 2006–2009. The SIAM core is a prior knowledge-based decision tree for MS

ABOUT THE AUTHOR

Andrea Baraldi (Laurea in Electronic Engineering, Univ. Bologna, 1989; Master in Software Engineering, Univ. Padua, 1994; PhD in Agricultural Sciences, Univ. Naples Federico II, 2017) has held research positions at the Italian Space Agency (2018–to date); Dept. Geoinformatics, Univ. Salzburg, Austria (2014–2017); Dept. Geographical Sciences, Univ. Maryland, College Park, MD (2010–2013); European Commission Joint Research Centre (2000–2002; 2005–2009); International Computer Science Institute, Berkeley, CA (1997–1999); European Space Agency Research Institute (1991–1993); Italian National Research Council (1989, 1994–1996, 2003–2004). In 2009, he founded Baraldi Consultancy in Remote Sensing. He was appointed with a Senior Scientist Fellowship at the German Aerospace Center, Oberpfaffenhofen, Germany (February 2014) and was a visiting scientist at the Ben Gurion Univ. of the Negev, Sde Boker, Israel (February 2015). His main interests center on the research and technological development of automatic near real-time spaceborne/airborne image pre-processing and understanding systems in operating mode, consistent with human visual perception. Dr. Baraldi served as an Associate Editor of the IEEE TRANS. NEURAL NETWORKS from 2001 to 2006.

PUBLIC INTEREST STATEMENT

Synonym of scene-from-image reconstruction and understanding, vision is an inherently ill-posed cognitive task; hence, it is difficult to solve and requires a priori knowledge in addition to sensory data to become better posed for numerical solution. In the inherently ill-posed cognitive domain of computer vision, this research was undertaken to validate by independent means a lightweight computer program for prior knowledge-based multi-spectral color naming, called Satellite Image Automatic Mapper (SIAM), eligible for automated near real-time transformation of large-scale Earth observation (EO) image datasets into European Space Agency (ESA) EO Level 2 information product, never accomplished to date at the ground segment. An original protocol for wall-to-wall thematic map quality assessment without sampling, where legends of the test and reference map pair differ and must be harmonized, was adopted. Conclusions are that SIAM is suitable for systematic ESA EO Level 2 product generation, regarded as necessary not sufficient pre-condition to transform EO big data into timely, comprehensive and operational EO value-adding information products and services.
reflectance space hyperpolyhedralization into static color names. Typically, a vocabulary of MS color names in a MS data (hyper)cube and a dictionary of land cover (LC) class names in the scene-domain do not coincide and must be harmonized (reconciled). The present Part 1—Theory provides the multidisciplinary background of a priori color naming. The subsequent Part 2—Validation accomplishes a GEO-CEOS stage 4 Val of the test SIAM-WELD annual map time-series in comparison with a reference 30 m resolution 16-class USGS National Land Cover Data 2006 map, based on an original protocol for wall-to-wall thematic map quality assessment without sampling, where the test and reference maps feature the same spatial resolution and spatial extent, but whose legends differ and must be harmonized.

Subjects: Algorithms & Complexity; Automation; Cognitive Artificial Intelligence; Expert systems; GIS, Remote Sensing & Cartography; Image Processing; Intelligent Systems; Machine Learning; Pattern Analysis; Quality Control & Reliability; Real-Time Systems; Systems Architecture

Keywords: Artificial intelligence; binary relationship; Cartesian product; cognitive science; color naming; connected-component multilevel image labeling; deductive inference; Earth observation; land cover taxonomy; high-level (attentive) and low-level (pre-attentional) vision; hybrid inference; image classification; image segmentation; inductive inference; machine learning-from-data; outcome and process quality indicators; radiometric calibration; remote sensing; surface reflectance; thematic map comparison; top-of-atmosphere reflectance; two-way contingency table; unsupervised data discretization/vector quantization; validation

1. Introduction

Jointly proposed by the intergovernmental Group on Earth Observations (GEO) and the Committee on Earth Observation Satellites (CEOS), the implementation plan for years 2005–2015 of the Global Earth Observation System of Systems (GEOSs) aimed at systematic transformation of multi-source Earth observation (EO) big data into timely, comprehensive and operational EO value-adding products and services (GEO, 2005), submitted to the GEO-CEOS Quality Assurance Framework for Earth Observation (QA4EO) calibration/validation (Cal/Val) requirements and suitable “to allow the access to the Right Information, in the Right Format, at the Right Time, to the Right People, to Make the Right Decisions” (Group on Earth Observation/Committee on Earth Observation Satellites (GEO-CEOS), 2010). In this definition of GEOSs, 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 big data challenges include capture, storage, search, sharing, transfer, analysis and visualization” (Wikipedia, 2018a), typically summarized as the five Vs of big data, specifically, volume, variety, velocity, veracity and value (IBM, 2016; Yang, Huang, Li, Liu, & Hu, 2017).

The GEOSs mission cannot be considered fulfilled by the remote sensing (RS) community to date. This is tantamount to saying the RS community is data-rich, but information-poor, a conjecture known as DRIP syndrome (Bernus & Noran, 2017). Before supporting this thesis with observations, the following definition is introduced. In this paper, an EO-IUS is defined in operating mode if and only if it scores “high” in every index of a minimally dependent and maximally informative (mDMI) set of EO outcome and process (OP) quantitative quality indicators (Q2Is), to be community-agreed upon to be used by members of the community, in agreement with the GEO-CEOS QA4EO Cal/Val guidelines (GEO-CEOS, 2010). A proposed instantiation of mDMI set of EO OP-Q2Is includes: (i) degree of automation, inversely related to human-machine interaction, (ii) effectiveness, e.g., thematic mapping accuracy, (iii) efficiency in computation time and in run-time memory occupation, e.g., inversely related to the number of system’s free-parameters to be user-defined based on heuristics, (iv) robustness (vice versa, sensitivity) to changes in input data, (v) robustness to changes in input parameters to be user-defined, (vi) scalability to changes in
user requirements and in sensor specifications, (vii) timeliness from data acquisition to information product generation, (viii) costs in manpower and computer power, (ix) value, e.g., semantic value of output products, economic value of output services, etc. (Baraldi, 2017, 2009; Baraldi & Boschetti, 2012a, 2012b; Baraldi, Boschetti, & Humber, 2014; Baraldi et al., 2010a, 2010b; Baraldi, Girona, & Simonetti, 2010c; Duke, 2016). According to the Pareto formal analysis of multi-objective optimization problems, optimization of an mDMI set of OP-Q^2Is is an inherently-ill posed problem in the Hadamard sense (Hadamard, 1902), where many Pareto optimal solutions lying on the Pareto efficient frontier can be considered equally good (Boschetti, Flassé, & Brivio, 2004). Any EO-IUS solution lying on the Pareto efficient frontier can be considered in operating mode, therefore suitable to cope with the five Vs of spatio-temporal EO big data (Yang et al., 2017). Several observations (true-facts) support this thesis. First, in 2012 the percentage of EO data ever downloaded from the European Space Agency (ESA) databases was estimated at about 10% or less (D’Elia 2012). This estimate is equal or superior (never inferior) to the percentage of ESA EO data ever used by the RS community. Since 2012, the same EO data exploitation indicator is expected to decrease, because any increase in productivity of existing EO-IUSs seems unable to match the exponential increase in the rate of collection of EO sensory data (NASA, 2016a). Second, EO-IUSs presented in the RS literature are typically assessed and compared based on the sole thematic mapping accuracy, which means their mDMI set of EO OP-Q^2Is remains largely unknown to date (Baraldi & Boschetti, 2012a, 2012b). As a consequence, the RS literature is unable to contradict the thesis that no EO-IUS is available in operating mode. For example, when EO data-derived thematic maps were generated by EO-IUSs based on a supervised (labeled) data learning approach, at continental or global spatial extent and with estimated accuracy not inferior to a target mapping accuracy requirement, the most limiting factors turned out to be the cost, timeliness, quality and availability of adequate supervised training data samples collected from field sites, existing maps or geospatial data archives in tabular form (Gutman et al., 2004). Third, no ESA EO data-derived Level 2 information product has ever been systematically generated at the ground segment (DLR and VEGA 2011; ESA 2015). In the ESA definition, an EO Level 2 information product is a single-date multi-spectral (MS) image radiometrically calibrated into surface reflectance (SURF) values corrected for atmospheric, adjacency and topographic effects, stacked with its data-derived scene classification map (SCM), whose legend includes quality layers cloud and cloud-shadow, starting from an ESA EO Level 1 product geometrically corrected and radiometrically calibrated into top-of-atmosphere reflectance (TOARF) values (European Space Agency (ESA), 2015; Deutsches Zentrum für Luft- und Raumfahrt e.V. (DLR) and VEGA Technologies, 2011; CNES, 2015).

This last observation deserves further discussion. In the words of Marr, “vision goes symbolic almost immediately without loss of information” (Marr, 1982). In agreement with this Marr’s intuition, ESA defines as EO Level 2 information product an information primitive (unit of information) consisting of a stack of two coupled (inter-dependent) variables, one sub-symbolic/numeric and one symbolic, where symbolic variable means both categorical and semantics. Equivalent to two sides of the same coin, these two variables are very closely related to each other and cannot be separated, even though they seem different. The first side of the ESA EO Level 2 information primitive is a multivariate numeric variable of the highest radiometric quality, related to the concept of quantitative/unequivocal information-as-thing in the terminology of philosophical hermeneutics (Capurro & Hjørland, 2003), see Figure 1. The second side of the ESA EO Level 2 information unit is an EO data-derived SCM, equivalent to a categorical variable of semantic value, related to the concept of qualitative/equivocal information-as-data-interpretation in
the terminology of philosophical hermeneutics (Capurro & Hjørland, 2003). In practice, ESA EO Level 2 product generation is a chicken-and-egg dilemma (Riano, Chuvieco, Salas, & Aguado, 2003), synonym of inherently ill-posed problem in the Hadamard sense (Hadamard, 1902). Therefore, it is very difficult to solve and requires a priori knowledge in addition to data to become better posed for numerical solution (Cherkassky & Mulier, 1998). On the one hand, no effective and efficient Cal of digital numbers (DNs) into SURF values corrected for atmospheric, topographic and adjacency effects is possible without an SCM, available a priori in addition to data to enforce a statistical stratification principle (Hunt & Tyrrell, 2012), synonym of layered (class-conditional) data analytics (Baraldi, 2017; Baraldi et al., 2010b; Baraldi & Humber, 2015; Baraldi, Humber, & Boschetti, 2013; Bishop & Colby, 2002; Bishop, Shroder, & Colby, 2003; DLR & VEGA, 2011; Dorigo, Richter, Bamler, & Wagner, 2009; Lück & van Niekerk, 2016; Riano et al., 2003; Richter & Schläpfer, 2012a, 2012b; Vermote & Saleous, 2007). On the other hand, no effective and efficient understanding (mapping) of a sub-symbolic EO image into a symbolic SCM is possible if DNs (pixels) are affected by low radiometric quality (GEO-CEOS, 2010). In an ESA EO Level 2 SCM product to be generated at the ground segment (midstream) as input to downstream applications, products and services (Mazzuccato & Robinson, 2017), the SCM legend is required to consist of a discrete, finite and hierarchical (multilevel) dictionary (Lipson, 2007; Mather, 1994; Swain & Davis, 1978) of general-purpose, user- and application-independent land cover (LC) classes, whose semantic value is “shallow” (not specialized) in hierarchy, but superior to zero semantics typical of numeric variables, in addition to quality layers cloud and cloud-shadow (ESA, 2015; DLR & VEGA, 2011; CNES, 2015). To our best knowledge, only one prototypical implementation of a sensor-specific ESA EO Level 2 product generator exists to date. Commissioned by ESA, the Sentinel 2 (atmospheric) Correction Prototype Processor (SEN2COR) is not run systematically at the ESA ground segment. Rather, it can be downloaded for free from the ESA web site to be run on user side (European Space Agency (ESA), 2015; Deutsches Zentrum für Luft- und Raumfahrt e.V. (DLR) and VEGA Technologies, 2011).

Noteworthy, a National Aeronautics and Space Administration (NASA) EO Level 2 product is defined as “a data-derived geophysical variable at the same resolution and location as Level 1 source data” (NASA, 2016b). Hence, dependence relationship “NASA EO Level 2 product → ESA EO Level 2 product” holds, where symbol “→” denotes relationship part-of from the supplier to the client, in agreement with the standard Unified Modeling Language (UML) for graphical modeling of object-oriented software (Fowler, 2003), see Figure 2. This dependence means that although space agencies and EO data distributors claim systematic NASA EO Level 2 product generation at the ground segment, this does not imply systematic ESA EO Level 2 product generation. Rather, the vice versa holds: if ESA EO Level 2 product generation is accomplished, then NASA EO Level 2 product generation is also fulfilled.

Different from the non-standard SEN2COR SCM legend instantiation (ESA, 2015; DLR & VEGA, 2011), one example of general-purpose, user- and application-independent ESA EO Level 2 SCM legend is the standard (community-agreed) 3-level 8-class dichotomous phase (DP) taxonomy of the Food and Agriculture Organization of the United Nations (FAO)—Land
Cover Classification System (LCCS) (Di Gregorio & Jansen, 2000). The FAO LCCS-DP hierarchy is “fully nested”. It comprises three dichotomous LC class-specific information layers, equivalent to a world ontology or world model (Di Gregorio & Jansen, 2000; Matsuyama & Hwang, 1990): DP Level 1—Vegetation versus non-vegetation, DP Level 2—Terrestrial versus aquatic and DP Level 3—Managed versus natural or semi-natural. The 3-level 8-class FAO LCCS-DP taxonomy is shown in Figure 3. For the sake of generality, a 3-level 8-class FAO LCCS-DP taxonomy is added with LC class “other”, synonym of “rest of the world” or “unknown”, which would include quality information layers cloud and cloud-shadow. In traditional EO image classification system design and implementation requirements (Swain & Davies, 1978), the presence of output class “unknown” is considered mandatory to cope with uncertainty in information-as-data-interpretation tasks. Hereafter, the standard 3-level 8-class FAO LCCS-DP taxonomy added with the mandatory output class “other”, which includes quality layers cloud and cloud-shadow, is identified as “augmented” 9-class FAO LCCS-DP taxonomy. In the complete two-phase FAO LCCS hierarchy, a general-purpose 3-level 8-class FAO LCCS-DP legend is preliminary to a high-level application-dependent and user-specific FAO LCCS Modular Hierarchical Phase (MHP) taxonomy, consisting of a hierarchical (deep) battery of one-class classifiers (Di Gregorio & Jansen, 2000), see Figure 3. In recent years, the two-phase FAO LCCS taxonomy has become increasingly popular (Ahlqvist, 2008). One reason of this popularity is that the FAO LCCS hierarchy is “fully nested” while alternative LC class hierarchies, such as the Coordination of Information on the Environment (CORINE) Land Cover (CLC) taxonomy (Bossard, Feranec, & Otahel, 2000), the U.S. Geological Survey (USGS) Land Cover Land Use (LCLU) taxonomy by J. Anderson (Lillesand & Kiefer, 1979), the International Global Biosphere Programme (IGBP) DISCover Data Set Land Cover Classification System (Belward, 1996) and the EO Image Librarian LC class legend (Dumitru, Cui, Schwarz, & Datcu, 2015), start from a Level 1 taxonomy which is already multi-class. In a hierarchical EO-IUS architecture submitted to a garbage in, garbage out (GIGO) information principle, synonym of error propagation through an information processing chain, the fully-nested two-phase FAO LCCS hierarchy makes explicit the full dependence of high-level EO OP-Q value features by lower-level LCCS-MHP data processing module, on previous EO OP-Q values featured by lower-level LCCS modules, starting from the initial FAO LCCS-DP Level 1 vegetation/non-vegetation information layer whose relevance in thematic mapping accuracy (vice versa, in error propagation) becomes paramount for all subsequent LCCS layers. The GIGO commonsense principle applied to hierarchical semantic dependence is neither trivial nor obvious to underline (Marcus, 2018). On the one hand, it agrees with a minor portion of the RS literature where supervised data learning classification of EO image datasets at continental or global spatial extent into binary LC class vegetation/non-vegetation is considered very challenging (Gutman et al., 2004). On the other hand, it is at odd with the RS mainstream, where the semantic information gap from sub-symbolic EO data to multi-class LC taxonomies is typically filled in one step, implemented as a supervised data learning classifier (Bishop, 1995; Cherkassky & Muller, 1998), e.g., a support vector machine, random forest or deep convolutional neural network (DCNN).
which is equivalent to an unstructured black box (Marcus, 2018), inherently semiautomatic and site specific (Liang, 2004) and whose opacity contradicts the well-known engineering principles of modularity, regularity and hierarchy typical of scalable systems (Lipson, 2007).

Starting from these premises our working hypothesis was that necessary not sufficient pre-condition for a yet-unfulfilled GEOSS development (GEO, 2005) is systematic generation at the ground segment of an ESA EO Level 2 product, never accomplished to date (ESA, 2015; DLR & VEGA, 2011), whose general-purpose SCM product is constrained as follows. First, the ESA EO Level 2 SCM legend agrees with the 3-level 9-class “augmented” FAO LCCS-DP taxonomy. Second, to comply with the GEO-CEOS QA4EO Cal/Val requirements, the SCM product must be submitted to a GEO-CEOS stage 4 Val, where an mDMI set of EO OP-Q2Is is evaluated by independent means at large spatial-extent and multiple time periods (GEO-CEOS, 2010). By definition, a GEO-CEOS stage 3 Val requires that “spatial and temporal consistency of the product with similar products are evaluated by independent means over multiple locations and time periods representing global conditions. In Stage 4 Val, results for Stage 3 are systematically updated when new product versions are released and as the time-series expands” (GEO-CEOS WGCV, 2015).

According to our working hypothesis, to contribute toward filling an analytic and pragmatic information gap from multi-source EO big data to ESA EO Level 2 information product as necessary not sufficient pre-condition to GEOSS development, the primary goal of this interdisciplinary study was to undertake an original (to the best of these authors’ knowledge, the first) outcome and process GEO-CEOS stage 4 Val of an off-the-shelf lightweight computer program, the Satellite Image Automatic Mapper™ (SIAM™), presented in recent years in the RS literature where enough information was provided for the implementation to be reproduced (Baraldi, 2017; Baraldi & Boschetti, 2012a, 2012b; Baraldi et al., 2010a, 2010b, 2010c; Baraldi & Humber, 2015; Baraldi et al., 2013; Baraldi, Puzzolo, Blonda, Brizzzone, & Tarantino, 2006, 2015; Baraldi, Tiede, Sudmanns, Belgui, & Lang, 2016). Implemented in operating mode in the C/C++ programming language, an off-the-shelf SIAM software executable runs: (i) automatically, i.e., it requires no human-machine interaction, (ii) in near real-time because it is non-iterative, more specifically it is one-pass, with a single subsystem which is two-pass (refer to the text below), and its computational complexity increases linearly with image size, and (iii)

(Cimpoi, Maji, Kokkinos, & Vedaldi, 2014), which is equivalent to an unstructured black box (Marcus, 2018), inherently semiautomatic and site specific (Liang, 2004) and whose opacity contradicts the well-known engineering principles of modularity, regularity and hierarchy typical of scalable systems (Lipson, 2007).

Starting from these premises our working hypothesis was that necessary not sufficient pre-condition for a yet-unfulfilled GEOSS development (GEO, 2005) is systematic generation at the ground segment of an ESA EO Level 2 product, never accomplished to date (ESA, 2015; DLR & VEGA, 2011), whose general-purpose SCM product is constrained as follows. First, the ESA EO Level 2 SCM legend agrees with the 3-level 9-class “augmented” FAO LCCS-DP taxonomy. Second, to comply with the GEO-CEOS QA4EO Cal/Val requirements, the SCM product must be submitted to a GEO-CEOS stage 4 Val, where an mDMI set of EO OP-Q2Is is evaluated by independent means at large spatial-extent and multiple time periods (GEO-CEOS, 2010). By definition, a GEO-CEOS stage 3 Val requires that “spatial and temporal consistency of the product with similar products are evaluated by independent means over multiple locations and time periods representing global conditions. In Stage 4 Val, results for Stage 3 are systematically updated when new product versions are released and as the time-series expands” (GEO-CEOS WGCV, 2015).

According to our working hypothesis, to contribute toward filling an analytic and pragmatic information gap from multi-source EO big data to ESA EO Level 2 information product as necessary not sufficient pre-condition to GEOSS development, the primary goal of this interdisciplinary study was to undertake an original (to the best of these authors’ knowledge, the first) outcome and process GEO-CEOS stage 4 Val of an off-the-shelf lightweight computer program, the Satellite Image Automatic Mapper™ (SIAM™), presented in recent years in the RS literature where enough information was provided for the implementation to be reproduced (Baraldi, 2017; Baraldi & Boschetti, 2012a, 2012b; Baraldi et al., 2010a, 2010b, 2010c; Baraldi & Humber, 2015; Baraldi et al., 2013; Baraldi, Puzzolo, Blonda, Brizzzone, & Tarantino, 2006, 2015; Baraldi, Tiede, Sudmanns, Belgui, & Lang, 2016). Implemented in operating mode in the C/C++ programming language, an off-the-shelf SIAM software executable runs: (i) automatically, i.e., it requires no human-machine interaction, (ii) in near real-time because it is non-iterative, more specifically it is one-pass, with a single subsystem which is two-pass (refer to the text below), and its computational complexity increases linearly with image size, and (iii)
in tile streaming mode, i.e., it requires a fixed run-time memory occupation. In addition to running on laptop and desktop computers, the SIAM lightweight computer program is eligible for use as mobile software application. By definition, a mobile software application is a lightweight computer program specifically designed to run on web services and/or mobile devices, such as tablet computers and smartphones, eventually provided with a mobile graphic user interface (GUI). An off-the-shelf SIAM software executable comprises six non-iterative subsystems for automated MS image analysis (decomposition) and synthesis (reconstruction) in linear time complexity. Its core is a one-pass prior knowledge-based decision tree (expert system) for MS reflectance space hyperpolyhedralization into static (non-adaptive-to-data) color names. Sketched in Figure 4, the SIAM software architecture is summarized as follows.

1. MS data radiometric calibration, in agreement with the GEO-CEOS QA4EO Cal requirements (GEO-CEOS, 2010). The SIAM expert system instantiates a physical data model; hence, it requires as input sensory data provided with a physical meaning. Specifically, DNs must be radiometrically Cal into a physical unit of radiometric measure to be community-agreed upon, such as TOARF values, SURF values or Kelvin degrees for thermal channels. Relationship TOARF $\supseteq$ SURF holds because SURF is a special case of TOARF in clear sky and flat terrain conditions (Chavez, 1988), i.e., TOARF $\approx$ SURF + atmospheric noise + topographic effects + surface adjacency effects. In a spectral decision tree for MS color space hyperpolyhedralization (partitioning), this relationship means that MS hyperpolyhedra (envelopes, manifolds) in “noisy” TOARF values include “noiseless” hyperpolyhedra in SURF values as special case of the former according to relationship subset-of, while the vice versa does not hold, see Figure 5.

2. One-pass prior knowledge-based SIAM decision tree for MS reflectance space hyperpolyhedralization into three static codebooks (vocabularies) of sub-symbolic/semi-symbolic color names as codewords, see Figure 5. Provided with inter-level parent–child relationships, the SIAM’s three-level vocabulary of static color names features a ColorVocabularyCardinality value which decreases from fine to intermediate to coarse, refer to Table 1 and Figure 6. MS reflectance space hyperpolyhedra for color naming are difficult to think of and impossible to visualize when the MS data space dimensionality is superior to three. This is not the case of basic color (BC) names adopted in human languages (Berlin & Kay, 1969), whose mutually exclusive and totally exhaustive perceptual polyhedra, neither necessarily convex nor connected, are intuitive to think of and easy to visualize in a 3D monitor-typical red-green-blue (RGB) data cube, see Figure 7 (Benavente, Vanrell, & Baldrich, 2008; Griffin, 2006). When each pixel of an MS image is mapped onto a color space partitioned into a set of mutually exclusive and totally exhaustive hyperpolyhedra equivalent to a vocabulary of BC names, then a 2D multilevel color map (2D
A gridded dataset of a multilevel variable is generated automatically (without human-machine interaction) in near real-time (with computational complexity increasing linearly with image size), where the number \( k \) of 2D map levels (color strata, color names) belongs to range \( \{1, \ldots, \text{ColorVocabularyCardinality}\} \). Popular synonyms of measurement space hyperpolyhedralization (discretization, partition) are vector quantization (VQ) in inductive machine learning-from-data (Cherkassky & Mulier, 1998; Elkan, 2003; Fritzke, 1997a, 1997b; Lee, Baeck, & Sung, 1997; Linde, Buzo, & Gray, 1980; Lloyd, 1982; Patanè and Russo, 2001, 2002), and deductive fuzzification of a numeric variable into fuzzy sets in fuzzy logic (Zadeh, 1965). Typical inductive learning-from-data VQ algorithms aim at minimizing a known VQ error function, e.g., a root mean square vector quantization error (RMSE), given a number of \( k \) discretization levels selected by a user based on \textit{a priori} knowledge and/or heuristic criteria. One of the most widely used VQ heuristics in RS and computer vision (CV) applications is the \( k \)-means VQ algorithm (Elkan, 2003; Lee et al., 1997; Linde et al., 1980; Lloyd, 1982), capable of convex Voronoi tessellation of a multi-variate data space (Cherkassky & Mulier, 1998; Fritzke, 1997a). For example, in a

### Table 1. The SIAM computer program is an EO system of systems scalable to any past, existing or future MS Imaging sensor provided with radiometric calibration metadata parameters

| SIAM, r88v7 | Input bands | Prior knowledge-based color map legends: Number of output spectral categories = Vocabulary of multi-spectral (MS) color names |
|-------------|-------------|-------------------------------------------------------------------------------------------------------------------|
|             | Fine        | Intermediate | Coarse | Inter-sensor discretization |
|             | discretization | discretization | discretization | levels (*) | levels (*) |
| L-SIAM      | 7—B, G, R, NIR, MIR1, MIR2, TIR | 96 | 48 | 18 | 33 (*); employed for inter-sensor post-classification change/no-change detection |
| S-SIAM      | 4—G, R, NIR, MIR1 | 68 | 40 | 15 |
| AV-SIAM     | 4—R, NIR, MIR1, TIR | 83 | 43 | 17 |
| Q-SIAM      | 4—B, G, R, NIR | 61 | 28 | 12 |

It encompasses the following subsystems. (i) 7-band Landsat-like SIAM™ (L-SIAM™), with input channels Blue (B), Green (G), Red (R), Near Infra-Red (NIR), Medium IR1 (MIR1), Medium IR2 (MIR2), and Thermal IR (TIR). (ii) 4-band (channels G, R, NIR, MIR1) SPOT-like SIAM™ (S-SIAM™). (iii) 4-band (channels R, NIR, MIR1, and TIR) Advanced Very High Resolution Radiometer (AVHRR)-like SIAM™ (AV-SIAM™). (iv) 4-band (channels B, G, R, and NIR) QuickBird-like SIAM™ (Q-SIAM™)
bag-of-words model applied to CV tasks, a numeric color space is typically discretized into a categorical color variable (codebook of codewords) by an inductive VQ algorithm, such as k-means; next, the categorical color variable is simplified by a 1st-order histogram representation, which disregards word grammar, semantics and even word-order, but keeps multiplicity; finally, the frequency of each color codeword is used as a feature for training a supervised data learning classifier (Cimpoi et al., 2014). Unlike the k-means VQ algorithm where the system’s free-parameter k is user-defined based on heuristics and the VQ error is estimated from the unlabeled dataset at hand, a user can fix the target VQ error value, so that it is the free-parameter k to be dynamically learned from the finite unlabeled dataset at hand by an inductive VQ algorithm (Patané & Russo, 2001, 2002), such as ISODATA (Memarsadeghi, Mount, Netanyahu, & Le Moigne, 2007). It means there is no universal number k of static hyperpolyhedra in a vector data space suitable for satisfying any VQ problem of interest if no target VQ error is specified in advance. As a viable strategy to cope with the inherent ill-posedness of inductive VQ problems (Cherkassky & Mulier, 1998), the SIAM expert system provides its three pre-
Figure 8. One segmentation map is deterministically generated from one multilevel image, such as a thematic map, but the vice versa does not hold, i.e., many multilevel images can generate the same segmentation map. In this example, nine image-objects/segments S1–S9 can be detected in the 3-level thematic map shown at left. Each segment consists of a connected set of pixels sharing the same multilevel map label. Each stratum/layer/level consists of one or more segments, e.g., stratum Vegetation (V) consists of two disjoint segments, S1 and S8. In any multilevel (categorical, nominal, qualitative) image domain, three labeled spatial primitives (spatial units) coexist and are provided with parent–child relationships: pixel with a level-label and a pixel identifier (ID, e.g., the row-column coordinate pair), segment (polygon) with a level-label and a segment ID, and stratum (multi-part polygon) with a level-label equivalent to a stratum ID. This overcomes the ill-fated dichotomy between traditional unlabeled sub-symbolic pixels versus labeled sub-symbolic segments in the numeric (quantitative) image domain traditionally coped with by the object-based image analysis (OBIA) paradigm (Blaschke et al., 2014).

defined VQ levels with a per-pixel RMSE estimation required for VQ quality assurance, in compliance with the GEO-CEOS QA4EO Val guidelines, refer to point (6) below.

(3) Well-posed (deterministic) two-pass detection of connected-components in the multilevel color map-domain (Dillencourt, Samet, & Tamminen, 1992; Sonka, Hlavac, & Boyle, 1994), where the number k of map levels belongs to range (1, ColorVocabularyCardinality), see Figure 8. These discrete and finite connected-components consist of connected sets of pixels featuring the same color label. Each connected-component is either (0D) pixel, (1D) line or (2D) polygon in the Open Geospatial Consortium (OGC) nomenclature (OGC 2015). They are typically known as superpixels in the CV literature (Achanta et al., 2011), homogeneous segments or image-objects in the object-based image analysis (OBIA) literature (Blaschke et al., 2014; Matsuyama & Hwang, 1990; Nagao & Matsuyama, 1980; Shackelford & Davis, 2003a, 2003b), and texture elements, i.e., texels, in human vision (Julesz, 1986; Julesz, Gilbert, Shepp, & Frisch, 1973). Whereas the physical model-based SIAM expert system requires no human-machine interaction to detect top-down superpixels whose shape and size can be any, superpixels detected bottom-up in statistical model-based CV algorithms typically require a pair of statistical model's free-parameters to be user-defined based on heuristics, such as a first heuristic-based geometric threshold equal to the superpixel maximum area and a second heuristic-based geometric threshold forcing a superpixel to stay compact in shape (Achanta et al., 2011). In a multilevel image domain where k is the number of levels (image-wide strata), individual (0D) pixels with label 1 to k, superpixels as connected sets of pixels featuring the same label 1 to k, and strata (layers), equal to discrete and finite collections of superpixels, mutually disjoint, but belonging to the same level 1 to k, co-exist as non-alternative labeled spatial units provided with a parent–child relationship, where each superpixel is a 2-tuple (superpixel ID, level 1-of-k) and each pixel is a 2-tuple (row-column coordinate pair, superpixel ID), see Figure 8.

(4) Well-posed 4- or 8-adjacency cross-aura representation in linear time of superpixel-contours, see Figure 9. These cross-aura contour values allow estimation of a scale-invariant planar shape index of compactness (Baraldi, 2017; Baraldi & Soares, 2017; Soares, Baraldi, & Jacobs, 2014), eligible for use by a high-level OBIA approach (Blaschke et al., 2014), see Figure 4.

(5) Superpixel/segment description table (Matsuyama & Hwang, 1990; Nagao & Matsuyama, 1980), to describe superpixels in a 1D tabular form (list) in combination with their 2D raster representation in the image-domain, referred to as “literal bit map” by Marr (1982), to take advantage of each data structure and overcome their shortcomings. Computationally, local spatial searches are more efficient in the 2D raster image-domain than in the 1D list representation, because “most of the spatial relationships that must be examined in early vision (encompassing the raw and full primal
sketch for token detection and texture segmentation, respectively) are rather local” (Marr, 1982).

Vice versa, if we had to examine global or “scattered, pepper-and-salt-like (spatial) configurations, then a (2D) bit map would probably be no more efficient than a (1D) list” (Marr, 1982).

(6) Superpixelwise-constant input image approximation (reconstruction), also known as “image-object mean view” in commercial OBIA applications (Trimble, 2015), followed by a per-pixel RMSE estimation between the original MS image and the reconstructed piecewise-constant MS image. This VQ error estimation strategy enforces a product quality assurance policy considered mandatory by the GEO-CEOS QA4EO Val guidelines. For example, VQ quality assurance supported by SIAM allows a user to adopt quantitative (objective) criteria in the selection of pre-defined VQ levels, equivalent to color names, to fit user- and application-specific VQ error requirement specifications.

An example of the SIAM output products automatically generated in linear time from a 13-band 10 m-resolution Sentinel-2A image radiometrically calibrated into TOARF values is shown in Figure 10.

The potential impact on the RS community of a GEO-CEOS stage 4 Val of an off-the-shelf SIAM lightweight computer program for automated near real-time prior knowledge-based MS reflectance space hyperpolyhedralization, superpixel detection and per-pixel VQ quality assessment is expected to be relevant, with special emphasis on existing or future hybrid (combined deductive and inductive) EO-IUSs. In the RS discipline, there is a long history of prior knowledge-based MS reflectance space partitioners for static color naming, alternative to SIAM’s, developed but never validated by space agencies, public organizations and private companies for use in hybrid EO-IUSs in operating mode, see Figure 11. Examples of hybrid EO image pre-processing applications in the quantitative/sub-symbolic domain of information-as-thing, where a numeric input variable is statistically class-conditioned (masked) by a static color naming first stage to generate as output another numeric variable considered more informative than the input one, are large-scale MS image compositing (Ackerman et al. 1998; Lück & van Niekerk, 2016; Luo, Trishchenko, & Khlopenkov, 2008), MS image atmospheric correction and topographic correction (Baraldi, 2017; Baraldi et al., 2010b; Baraldi & Humber, 2015; Baraldi et al., 2013; Bishop & Colby, 2002; Bishop et al., 2003; DLR & VEGA, 2011; Dorigo et al., 2009; Lück & van Niekerk, 2016; Riano et al., 2003; Richter & Schlüpf, 2012a, 2012b; Vermote & Saleous, 2007), see Figure 12, MS image adjacency effect correction (DLR & VEGA, 2011) and radiometric quality assessment of pan-sharpened MS imagery (Baraldi, 2017; Despini, Teggi, & Baraldi, 2014). Examples of hybrid EO image classification applications in the qualitative/equivocal/categorical domain of information-as-data-interpretation and statistically class-conditioned by a static color naming first stage are cloud and cloud-shadow quality layer detection (Baraldi, 2015, 2017; Baraldi et al., 2011; Lück & van Niekerk, 2016), single-date LC classification (DLR & VEGA, 2011; GeoTerraImage, 2015; Lück & van Niekerk, 2016; Muirhead & Malkawi, 1989; Simonetti et al. 2015a), multi-temporal post-classification LC change (LCC)/no-change detection (Baraldi, 2017; Baraldi et al., 2016; Simonetti et al., 2015a; Tiede,
Baraldi, Sudmanns, Belgiu, & Lang, 2016, multi-temporal vegetation gradient detection and quantization into fuzzy sets (Arvor, Madiela, & Corpetti, 2016), multi-temporal burned area detection (Boschetti, Roy, Justice, & Humber, 2015), and prior knowledge-based LC mask refinement (cleaning) of supervised data samples employed as input to supervised data learning EO-IUSs (Baraldi et al., 2010a, 2010b). Due to their large application domain, ranging from low- (pre-attentional) to high-level (attentional) vision tasks, existing hybrid EO-IUSs in operating mode, whose statistical data models are class-conditioned by static color naming, become natural candidates for the research and development (R&D) of an EO-IUS in operating mode, capable of systematic transformation of multi-source single-date MS imagery into ESA EO Level 2 product at the ground segment.

Figure 10. See note.
Figure 11. Same as in Schläpfer et al. (2009), courtesy of Daniel Schläpfer, ReSe Applications Schläpfer. A complete ("augmented") hybrid inference workflow for MS image correction from atmospheric, adjacency and topographic effects. It combines a standard Atmospheric/Topographic Correction for Satellite Imagery (ATCOR) commercial software workflow (Richter & Schläpfer, 2012a, 2012b), with a bidirectional reflectance distribution function (BRDF) effect correction. Processing blocks are represented as circles and output products as rectangles. This hybrid (combined deductive and inductive) workflow alternates deductive/prior knowledge-based with inductive/learning-from-data inference units, starting from initial conditions provided by a first-stage deductive Spectral Classification of surface reflectance signatures (SPECL) decision tree for color naming (pre-classification), implemented within the ATCOR commercial software toolbox (Richter & Schläpfer, 2012a, 2012b). Categorical variables generated by the pre-classification and classification blocks are employed to stratify (mask) unconditional numeric variable distributions, in line with the statistic stratification principle (Hunt & Tyrrell, 2012). Through statistic stratification (class-conditional data analytics), inherently ill-posed inductive learning-from-data algorithms are provided with prior knowledge required in addition to data to become better posed for numerical solution, in agreement with the machine learning-from-data literature (Cherkassky & Mulier, 1998).
The terminology adopted in the rest of this paper is mainly driven from the multidisciplinary domain of cognitive science, see Figure 13. Popular synonyms of deductive inference are top-down, prior knowledge-based, learning-from-rule and physical model-based inference. Synonyms of inductive inference are bottom-up, learning-from-data, learning-from-examples and statistical model-based inference (Baraldi, 2017; Baraldi & Boschetti, 2012a, 2012b; Liang, 2004). Hybrid inference systems combine statistical and physical data models to take advantage of the unique features of each and overcome their shortcomings (Baraldi, 2017; Baraldi & Boschetti, 2012a, 2012b; Cherkassky & Mulier, 1998; Liang, 2004). For example, in biological cognitive systems “there is never an absolute beginning” (Piaget, 1970), where an a priori genotype provides initial conditions to an inductive learning-from-examples phenotype (Parisi, 1991). Hence, any biological cognitive system is a hybrid inference system where inductive/phenotypic learning-from-examples mechanisms explore the neighborhood of deductive/genotypic initial conditions in a solution space (Parisi, 1991). In line with biological cognitive systems, an artificial hybrid inference system can alternate deductive and inductive inference algorithms, starting from a deductive inference first stage for initialization purposes, see Figure 11. It means that no deductive inference subsystem, such as SIAM, should be considered stand-alone, but eligible for use in a hybrid inference system architecture to initialize (pre-condition, stratify) inductive learning-from-examples algorithms, which are inherently ill-posed, difficult to solve and require a priori knowledge in addition to data to become better posed for numerical solution, as clearly acknowledged by the machine learning-from-data literature (Bishop, 1995; Cherkassky & Mulier, 1998).

To comply with the GEO-CEOS stage 4 Cal/Val requirements, the selected ready-for-use SIAM software executable had to be validated by independent means on a radiometrically calibrated EO image time-series at large spatial extent. This input data set was identified in the open-access USGS 30 m resolution Web Enabled Landsat Data (WELD) annual composites of the conterminous U.S. (CONUS) for the years 2006–2009, radiometrically calibrated into TOARF values (Homer, Huang, Yang, Wylie, & Coan, 2004; Roy et al., 2010; WELD 2015). The 30 m resolution 16-class U.S. National Land Cover Data (NLCD) 2006 map, delivered in 2011 by the USGS Earth Resources Observation Systems (EROS) Data Center (EDC) (EPA 2007; Vogelmann et al., 2001; Vogelmann, Sohl, Campbell, & Shaw, 1998; Wickham, Stehman, Fry, Smith, & Homer, 2010; Wickham et al., 2013; Xian & Homer, 2010), was selected as reference thematic map at continental spatial extent. The USGS 16-class NLCD 2006 map legend is summarized in Table 2. To account for typical non-stationary geospatial statistics, the USGS NLCD 2006 thematic map was partitioned into 86 Level III ecoregions of North America collected from the Environmental Protection Agency (EPA) (EPA 2013; Griffith & Omernik, 2009).

In this experimental framework, the test SIAM-WELD annual color map time-series for the years 2006–2009 and the reference USGS NLCD 2006 map share the same spatial extent and spatial resolution, but their map legends are not the same. These working hypotheses are neither trivial nor conventional in the RS literature, where thematic map quality assessment strategies typically adopt an either random or non-random sampling strategy and assume that the test and reference
thematic map dictionaries coincide (Stehman & Czaplewski, 1998). Starting from a stratified random sampling protocol presented in Baraldi et al. (2014), the secondary contribution of the present study was to develop a novel protocol for wall-to-wall comparison without sampling of two thematic maps featuring the same spatial extent and spatial resolution, but whose legends can differ.

For the sake of readability this paper is split into two, the present Part 1—Theory and the subsequent Part 2—Validation. An expert reader familiar with static color naming in cognitive science, spanning from linguistics to human vision and CV, can skip the present Part 1, either totally or in part. To make this paper self-contained and provided with a relevant survey value, the Part 1 is organized as follows. The multidisciplinary background of color naming is discussed in Chapter 2. Chapter 3 reviews the long history of prior knowledge-based decision trees for MS color naming presented in the RS literature. To cope with thematic map legends that do not coincide and must be harmonized (reconciled, associated, translated) (Ahlqvist, 2005), such as dictionaries of MS color names in the image-domain and LC class names in the scene-domain, Chapter 3 proposes an original hybrid inference guideline to identify a categorical variable-pair relationship, where prior beliefs are combined with additional evidence inferred from new data. An original measure of categorical variable-pair association (harmonization) in a binary relationship is proposed in Chapter 4. In the subsequent Part 2, GEO-CEOS stage 4 Val results are collected by an original protocol for wall-to-wall thematic map quality assessment without sampling, where legends of the test SIAM-WELD annual map time-series and reference USGS NLCD 2006 map are harmonized. Conclusions are that the annual SIAM-WE LD map time-series for the years 2006–2009 provides a first example of GEO-CEOS stage 4 validated ESA EO Level 2 SCM product, where the Level 2 SCM legend is the “augmented” 2-level 4-class FAO LCCS taxonomy at the DP Level 1 (vegetation/non-vegetation) and DP Level 2 (terrestrial/aquatic), added with extra class “rest of the world”.

2. Problem background of color naming in cognitive science
Within the cognitive science domain, vision is synonym of scene-from-image reconstruction and understanding, see Figure 13. Encompassing both biological vision and CV, vision is a cognitive (information-as-data-interpretation) problem inherently ill-posed in the Hadamard sense (Hadamard, 1902); hence, it is very difficult to solve. Vision is non-polynomial (NP)-hard in computational complexity (Frintrop, 2011; Tsotsos, 1990) and requires a priori knowledge in addition to sensory data to become better posed for numerical solution (Cherkassky & Mulier, 1998). It is inherently ill-posed because affected by, first, data dimensionality reduction from the 4D spatio-temporal scene-domain to the (2D) image-domain and, second, by a semantic information gap from ever-varying sensations in
| Code | ID | Name                      | Land cover (LC) Class Definition                                                                 | ID         |
|------|----|---------------------------|--------------------------------------------------------------------------------------------------|------------|
| 11   | OW | Open water                | OW: Areas of open water, generally with less than 25% cover of vegetation or soil                | B4 — Non-vegetated aquatic |
| 12   | PIS | Perennial Ice/Snow        | PIS: Areas characterized by a perennial cover of ice and/or snow, generally greater than 25% of total cover. | B4         |
| 21   | DOS | Developed, Open Space     | DOS: Includes areas with a mixture of some constructed materials, but mostly vegetation in the form of lawn grasses. Impervious surfaces account for less than 20 percent of total cover. These areas most commonly include large-lot single-family housing units, parks, golf courses, and vegetation planted in developed settings for recreation, erosion control, or aesthetic purposes. DLI, DMI, DHI: refer to the “National Land Cover Database 2006 (NLCD2006),” Multi-Resolution Land Characteristics Consortium (MRLC), 2013. | B3 — Non-vegetated terrestrial/A1 — Vegetated terrestrial |
| 22   | DLI | Developed, Low Intensity  |                                                                                                 |            |
| 23   | DMI | Developed, Medium Intensity |                                                                                                 |            |
| 24   | DHI | Developed, High Intensity |                                                                                                 |            |
| 31   | BL  | Barren Land (Rock/Sand/Clay) | BL: Barren areas of bedrock, desert pavement, scarps, talus, slides, volcanic material, glacial debris, sand dunes, strip mines, gravel pits and other accumulations of earthen material. Generally, vegetation accounts for less than 15% of total cover. As a consequence of this constraint, class BL covers only 1.21% of the CONUS total surface. | B3         |
| 41   | DF  | Deciduous Forest          | DF: Areas dominated by trees generally greater than 5 m tall, and greater than 20% of total vegetation cover. More than 75 percent of the tree species shed foliage simultaneously in response to seasonal change. | A1         |
| 42   | EF  | Evergreen Forest          | EF: Areas dominated by trees generally greater than 5 m tall, and greater than 20% of total vegetation cover. More than 75 percent of the tree species maintain their leaves all year. Canopy is never without green foliage. |            |
| 43   | MF  | Mixed Forest              | MF: Mixed Forest—Areas dominated by trees generally greater than 5 m tall, and greater than 20% of total vegetation cover. Neither deciduous nor evergreen species are greater than 75 percent of total tree cover. |            |

(Continued)
Table 2. (Continued)

USGS NLCD 2001/2006/2011 Classification Scheme (Legend), Level II & FAO LCCS-DP, level 1: A = Veg, B = Non-Veg, and level 2: 1 = Terrestrial, 2 = Aquatic

| Code | ID      | Name          | Land cover (LC) Class Definition                                                                                                                                                                                                 | ID      |
|------|---------|---------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------|
| 51   | -       | Dwarf Scrub   | SS: Areas dominated by shrubs; less than 5 m tall with shrub canopy typically greater than 20% of total vegetation. This class includes true shrubs, young trees in an early successional stage or trees stunted from environmental conditions. The aforementioned definition of class BL means that class SS may feature a vegetated cover which accounts for 15% of total cover or more. | A1/B3   |
| 52   | SS      | Scrub/Shrub   |                                                                                                               |         |
| 71   | GH      | Grassland/Herbaceous Sedge | GH: Areas dominated by grammanoid or herbaceous vegetation, generally greater than 80% of total vegetation. These areas are not subject to intensive management such as tilling, but can be utilized for grazing. The aforementioned definition of class BL means that class GH may feature a vegetated cover which accounts for 15% of total cover or more. | A1/B3   |
| 72   | -       | Grassland/Herbaceous Lichens |                                                                                                               |         |
| 73   | -       | Grassland/Herbaceous |                                                                                                               |         |
| 74   | -       | Grassland/Herbaceous Moss |                                                                                                               |         |
| 81   | PH      | Pasture/Hay   | PH: Areas of grasses, legumes, or grass-legume mixtures planted for livestock grazing or the production of seed or hay crops, typically on a perennial cycle. Pasture/hay vegetation accounts for greater than 20 percent of total vegetation. | A1      |
| 82   | CC      | Cultivated Crops |                                                                                                               |         |
| 90   | WW      | Woody Wetlands | WW: Areas where forest or shrubland vegetation accounts for greater than 20 percent of vegetation cover and the soil or substrate is periodically saturated with or covered with water. EHW: Areas where perennial herbaceous vegetation accounts for greater than 80% of vegetative cover and the soil or substrate is periodically saturated with or covered with water. | A2      |
| 95   | EHW     | Emergent Herbaceous Wetland |                                                                                                               |         |

For further details, refer to the “National Land Cover Database 2006 (NLCD2006),” Multi-Resolution Land Characteristics Consortium (MRLC), 2013. The right column instantiates a possible binary relationship R: A ⇒ B ⊆ A × B from set A = NLCD legend to set B = 2-level 4-class Dichotomous Phase (DP) taxonomy of the Food and Agriculture Organization of the United Nations (FAO)—Land Cover Classification System (LCCS) (Di Gregorio & Jansen, 2000), refer to Figure 3.
the (2D) image-domain to stable percepts in the mental model of the 4D scene-domain (Matsuyama & Hwang, 1990). On the one hand, ever-varying sensations collected from the 4D spatio-temporal physical world are synonym of observables, numeric/quantitative variables of sub-symbolic value or sensory data provided with a physical unit of measure, such as TOARF or SURF values, but featuring no semantics corresponding to abstract concepts, like perceptual categories or mental states. On the other hand, in a modeled world, also known as world ontology, mental world or “world model” (Matsuyama & Hwang, 1990), stable percepts are nominal/categorical/qualitative variables of symbolic value, i.e., they are categorical variables provided with semantics, such as LC class names belonging to a hierarchical FAO LCCS taxonomy of the world (Di Gregorio & Jansen, 2000), see Figure 3.

In statistics, the popular concept of latent/hidden variables was introduced to fill the information gap from input observables to target categorical variables. Latent/hidden variables are not directly measured, but inferred from observable numeric variables to link sensory data in the real world to categorical variables of semantic quality in the modeled world. “The terms hypothetical variable or hypothetical construct may be used when latent variables correspond to abstract concepts, like perceptual categories or mental states” (Baraldi, 2017; Shotton, Winn, Rother, & Criminisi, 2009; Wikipedia, 2018b). Hence, to fill the semantic gap from low-level numeric variables of sub-symbolic quality to high-level categorical variables of semantic value, hypothetical variables, such as categorical BC names (Benavente et al., 2008; Berlin & Kay, 1969; Gevers, Gijsenij, van de Weijer, & Geusebroek, 2012; Griffin, 2006), are expected to be mid-level categorical variables of “semi-symbolic” quality, i.e., hypothetical variables are nominal variables provided with a semantic value located “low” in a hierarchical ontology of the world, such as the hierarchical FAO LCCS taxonomy (Di Gregorio & Jansen, 2000), but always superior to zero, where zero is the semantic value of sub-symbolic numeric variables, see Figure 14.

In vision, spatial topological and spatial non-topological information components typically dominate color information (Baraldi, 2017; Matsuyama & Hwang, 1990). This thesis is proved by the undisputable fact that achromatic (panchromatic) human vision, familiar to everybody when wearing sunglasses, is nearly as effective as chromatic vision in scene-from-image reconstruction and understanding. Driven from perceptual evidence in human vision typically investigated by cognitive science, see Figure 13, a necessary not sufficient condition for a CV system to prove it fully exploits spatial topological and spatial non-topological information components in addition to color is to perform nearly the same when input with either panchromatic or color imagery. Stemming from a priori knowledge of human vision available in addition to sensory data, this necessary not sufficient condition can be adopted to make an inherently ill-posed CV system design and implementation problem better constrained for numerical solution.

Deeply investigated in CV (Frintrop, 2011; Sonka et al., 1994), content-based image retrieval (Smeulders, Worring, Santini, Gupta, & Jain, 2000) and EO image applications proposed by the RS community (Baraldi, 2017; Baraldi & Boschetti, 2012a, 2012b; Matsuyama & Hwang, 1990; Nagao & Matsuyama, 1980; Shackelford & Davis, 2003a, 2003b), well-known visual features are: (i) color values, typically discretized by humans into a finite and discrete vocabulary of BC...
Figure 15. Example of 1D image analysis. Synonym of 1D analysis of a 2D gridded dataset, it is affected by spatial data dimensionality reduction. The (2D) image at left is transformed into the 1D vector data stream shown at bottom, where vector data are either pixel-based or spatial context-sensitive, e.g., local window-based. This 1D vector data stream, either pixel-based or local window-based, means nothing to a human photointerpreter. When it is input to a traditional inductive data learning classifier, this 1D vector data stream is what the inductive classifier actually sees when watching the (2D) image at left. Undoubtedly, computers are more successful than humans in 1D image analysis, invariant to permutations in the input vector data sequence, such as in orderless pooling encoders (Cimpoi et al., 2014). Nonetheless, humans are still far more successful than computers in 2D image analysis, synonym of spatial topology-preserving (retinotopic) image analysis (Tsotsos, 1990), sensitive to permutations in the input vector data sequence, such as in order-sensitive pooling encoders (Cimpoi et al., 2014).

Neglecting the fact that spatial topological and spatial non-topological information components typically dominate color information in both the (2D) image-domain and the 4D spatio-temporal scene-domain involved with the cognitive task of vision (Matsuyama & Hwang, 1990), traditional EO-IUSs adopt a 1D image analysis approach, see Figure 15. In 1D image analysis, a 1D streamline of vector data, either spatial context-sensitive (e.g., window-based or image object-based like in OBIA approaches) or spatial context-insensitive (pixel-based), is processed insensitive to changes in the order of presentation of the input sequence. In practice 1D image analysis is invariant to permutations, such as in orderless pooling encoders (Cimpoi et al., 2014). When vector data are spatial context-sensitive then 1D image analysis ignores spatial topological information. When vector data are pixel-based then 1D image analysis ignores both spatial topological and spatial non-topological information components. Prior knowledge-based color naming of a spatial unit \( x \) in the image-domain, where \( x \) is either (0D) point, (1D) line or (2D) polygon defined according to the OG C nomenclature (OGC, 2015), is a special case of 1D image analysis, either pixel-based or image object-based, where spatial topological and/or spatial non-topological information are ignored, such as in SIAM’s static color naming (Baraldi et al., 2006).

Alternative to 1D image analysis, 2D image analysis relies on a sparse (distributed) 2D array (2D regular grid) of local spatial filters, suitable for spatial topology-preserving (retinotopic) feature mapping (DiCarlo, 2017; Fritzke, 1997a; Martineitz, Berkovich, & Schulten, 1994; Tsotsos, 1990), sensitive to permutations in the input vector data sequence, such as in order-sensitive pooling encoders (Cimpoi et al., 2014), see Figure 16. The human brain’s organizing principle is topology-preserving feature mapping (Feldman, 2013). In the biological visual system, topology-preserving feature maps are primarily spatial, where activation domains of physically adjacent processing units in the 2D array of convolutional filters are spatially adjacent regions in the 2D visual field. Provided with a superior degree of biological plausibility in modeling 2D spatial topological and spatial non-topological information components, distributed processing systems capable of 2D image analysis, such as physical model-based (“hand-crafted”) 2D wavelet filter banks (Mallat, 2016) and end-to-end inductive learning-from-data DCNNs, typically
outperform 1D image analysis approaches (Cimpoi et al., 2014; DiCarlo, 2017), although DCNNs are the subject of increasing criticisms by the artificial intelligence (AI) community (DiCarlo, 2017; Etzioni, 2017; Marcus, 2018). This apparently trivial consideration is at odd with a relevant portion of the RS literature, where pixel-based 1D image analysis is mainstream, followed in popularity by spatial context-sensitive 1D image analysis implemented within the OBIA paradigm (Blaschke et al., 2014). Undoubtedly, computers are more successful than humans in 1D image analysis, invariant to permutations in the input vector data sequence (Cimpoi et al., 2014). Nonetheless, humans are still far more successful than computers in 2D image analysis, synonym of spatial topology-preserving feature mapping (Tsotsos, 1990), which implies sensitivity to permutations in the input vector data sequence (Cimpoi et al., 2014).

Since traditional EO-IUSs adopt a 1D image analysis approach, where dominant spatial information is omitted either totally or in part in favor of secondary color information, it is useful to turn attention to the multidisciplinary framework of cognitive science to shed light on how humans cope with color information. According to cognitive science which includes linguistics, the study of languages, see Figure 13, humans discretize (fuzzify) ever-varying quantitative (numeric) photometric and spatio-temporal sensations into stable qualitative/categorical/nominal percepts, eligible for use in symbolic human reasoning based on a convergence-of-evidence approach (Matsuyama & Hwang, 1990). In their seminal work, Berlin and Kay proved that 20 human languages, spoken across space and time in the real world, partition quantitative color sensations collected in the visible portion of the electromagnetic spectrum, see Figure 1, onto the same “universal” vocabulary of eleven BC names (Berlin & Kay, 1969): black, white, gray, red, orange, yellow, green, blue, purple, pink and brown. In a 3D monitor-typical red-green-blue (RGB) data cube, BC names are intuitive to think of and easy to visualize. They provide a mutually exclusive and totally exhaustive partition of a monitor-typical RGB data cube into RGB polyhedra neither necessarily connected nor convex, see Figure 7 (Benavente et al., 2008; Griffin, 2006). Since they are community-agreed upon to be used by members of the same community, RGB BC polyhedra are prior knowledge-based, i.e., stereotyped, non-adaptive-to-data (static), general-purpose, application- and user-independent. Multivariate measurement space partitioning into a discrete and finite set of mutually exclusive and totally exhaustive hyperpolyhedra copes with the transformation of a numeric variable into a categorical variable, see Figure 7. Numeric variable discretization is a typical problem in many scientific disciplines, such as inductive VQ in machine learning-from-data (Cherkassky & Mulier, 1998) and
deductive numeric variable fuzzification into discrete fuzzy sets, e.g., low, medium and high, in fuzzy logic (Zadeh, 1965), refer to Chapter 1.

To summarize, human languages refer to human colorimetric perception in terms of a stable, prior knowledge-based vocabulary (codebook) of BC names (codewords) non-adaptive to data, physically equivalent to a discrete and finite set of mutually exclusive and totally exhaustive hyperpolyhedra, neither necessarily convex nor connected in a numeric MS color space, identified as $\mathfrak{m}^{\text{MS}}$, where $\text{MS} \geq 2$, e.g., $\text{MS} = 3$ like in a monitor-typical RGB data cube, see Figure 7. These BC names are conceptually equivalent to a latent/hypothetical categorical variable of semi-symbolic quality, see Figure 14, capable of linking sub-symbolic sensory data in the real world, specifically color values in color space $\mathfrak{m}^{\text{MS}}$, to categorical variables of semantic (symbolic) quality in the world model, also known as world ontology or mental world, made of abstract concepts, like perceptual categories of real-world objects or mental states.

In an analytic model of vision based on a convergence-of-evidence approach, the first original contribution of the present Part 1 is to encode prior knowledge about color naming into a CV system by design, as described hereafter. Irrespective of their Pearson inter-feature cross-correlation, if any, it is easy to prove that individual sources of visual evidence, such as color, local shape, texture and inter-object spatial relationships, are statistically independent because, in general, Pearson's linear cross-correlation does not imply causation (Baraldi, 2017; Baraldi & Soares, 2017; Pearl, 2009). According to a “naive” hypothesis of conditional independence of visual features color, local shape, texture and inter-object spatial relationships, when target classes of observed objects in the real-world scene are $c = 1, \ldots, \text{ObjectClassLegendCardinality}$, for a given discrete spatial unit $x$ in the image-domain, either 0D point, 1D line or 2D polygon (OGC, 2015), then the well-known “naive” Bayes classification formulation (Bishop, 1995) becomes

$$p(c|\text{ColorValue}(x), \text{ShapeValue}(x), \text{TextureValue}(x), \text{SpatialRelationships}(x, \text{Neigh}(x)))$$

$$= p(c|F_1, \ldots, F_{I-4}) = p(c) \prod_{i=1}^{I-4} p(F_i|c) = p(c) \cdot p(\text{ColorValue}(x)|c) \cdot p(\text{ShapeValue}(x)|c)$$

$$\cdot p(\text{TextureValue}(x)|c) \cdot p(\text{SpatialRelationships}(x, \text{Neigh}(x))|c) \leq \min\{p(c|\text{ColorValue}(x)), p(c|\text{ShapeValue}(x)), p(c|\text{TextureValue}(x)), p(c|\text{SpatialRelationships}(x, \text{Neigh}(x))))\}, c = 1, \ldots, \text{ObjectClassLegendCardinality}$$

(1)

where $\text{ColorValue}(x)$ belongs to a MS measurement space $\mathfrak{m}^{\text{MS}}$, i.e., $\text{ColorValue}(x) \in \mathfrak{m}^{\text{MS}}$, and $\text{Neigh}(x)$ is a generic 2D spatial neighborhood of spatial unit $x$ in the (2D) image-domain. Equation (1) shows that any convergence-of-evidence approach is more selective than each individual source of evidence, in line with a focus-of-visual attention mechanism (Frintrop, 2011). For the sake of simplicity, if priors are ignored because considered equiprobable in a maximum class-conditional likelihood inference approach alternative to a maximum a posteriori optimization criterion, then Equation (1) becomes

$$p(c|\text{ColorValue}(x), \text{ShapeValue}(x), \text{TextureValue}(x), \text{SpatialRelationships}(x, \text{Neigh}(x))) \propto$$

$$p(\text{ColorValue}(x)|c) \cdot p(\text{ShapeValue}(x)|c) \cdot p(\text{TextureValue}(x)|c) \cdot p(\text{SpatialRelationships}(x, \text{Neigh}(x))|c)$$

$$\cdot \prod_{i=1}^{\text{ColorVocabularyCardinality}} p(\text{ColorValue}(x)|\text{ColorName}) p(\text{ColorName}|c)$$

$$= 1, \ldots, \text{ObjectClassLegendCardinality}$$

(2)

where color space $\mathfrak{m}^{\text{MS}}$ is partitioned into hyperpolyhedra, equivalent to a discrete and finite vocabulary of static color names, with $\text{ColorName} = 1, \ldots, \text{ColorVocabularyCardinality}$. To further simplify Equation (2), its canonical interpretation based on frequentist statistics can be relaxed by fuzzy logic (Zadeh, 1965), so that the logical-AND operator is replaced by a fuzzy-AND (min) operator, inductive class-conditional probability $p(x|c) \in [0, 1]$, where
\begin{align*}
\sum_{c=1}^{ObjectClassLegendCardinality} p(x|c) &= 0, \text{ is replaced by a deductive membership (compatibility) function } m(x|c) \geq 0 \text{ according to the principles of fuzzy logic, where compatibility/membership does not mean probability, and color space hyperpolyhedra are considered mutually exclusive and totally exhaustive. If these simplifications are adopted, then Equation (2) becomes} \\
&\quad m(c|\text{ColorValue}(x), \text{ShapeValue}(x), \text{TextureValue}(x), \text{SpatialRelationships}(x, \text{Neigh}(x))) \propto \min \\
&\quad \left( \sum_{\text{ColorName}=1}^{\text{ColorVocabularyCardinality}} m(\text{ColorValue}(x)|\text{ColorName}) m(\text{ColorName}|c), m(\text{ShapeValue}(x)| c), m(\text{TextureValue}(x)| c), m(\text{SpatialRelationships}(x, \text{Neigh}(x))| c) \right), c = 1, \ldots, \text{ObjectClassLegendCardinality}, \text{ where ColorName}^* \in \{1, \text{ColorVocabularyCardinality}\}, \text{ such that } m(\text{ColorValue}(x)|\text{ColorName}^*) = 1 \text{ and } m(\text{ColorName}^*|c) \in \{0, 1\}.
\end{align*}

In Equation (3), the following considerations hold.

- Each numeric ColorValue(x) in color space \( \mathbb{Y}^{MS} \) belongs to a single color name (hyperpolyhedron) ColorName* in the static color name vocabulary, i.e., \( \forall \text{ColorValue}(x) \in \mathbb{Y}^{MS}, \sum_{\text{ColorName}=1}^{\text{ColorVocabularyCardinality}} m(\text{ColorValue}(x)|\text{ColorName}) = m(\text{ColorValue}(x)|\text{ColorName}^*) = 1 \) holds, where \( m(\text{ColorValue}(x)|\text{ColorName}) \in \{0, 1\} \) is a binary (crisp, hard) membership function, with ColorName = 1, ..., ColorVocabularyCardinality and ColorName* \in \{1, \text{ColorVocabularyCardinality}\}.

- Set \( A = \text{VocabularyOfColorNames} \), with cardinality \( |A| = a = \text{ColorVocabularyCardinality} \), and set \( B = \text{LegendOfObjectClassNames} \), with cardinality \( |B| = b = \text{ObjectClassLegendCardinality} \), can be considered a bivariate categorical random variable where two univariate categorical variables A and B are generated from a single population. A binary relationship from set A to set B, R: \( A \rightarrow B \), is a subset of the 2-fold Cartesian product (product set) \( A \times B \), whose size is \( \text{rows} \times \text{columns} = a \times b \), hence, \( R: A \rightarrow B \subseteq A \times B \). The Cartesian product of two sets \( A \times B \) is a set whose elements are ordered pairs. Hence, the Cartesian product is non-commutative, \( A \times B \neq B \times A \). In agreement with common sense, see Table 3, binary relationship R: VocabularyOfColorNames \( \Rightarrow \) LegendOfObjectClassNames is a set of ordered pairs where each ColorName can be assigned to none, one or several classes of observed scene-objects with class index \( c = 1, \ldots, \text{ObjectClassLegendCardinality} \), whereas each class of observed objects can be assigned with none, one or several color names to define the class-specific colorimetric attribute. Binary membership values \( m(\text{ColorName}|c) \in \{0, 1\} \) and \( m(c|\text{ColorName}) \in \{0, 1\} \), with \( c = 1, \ldots, \text{ObjectClassLegendCardinality} \) and ColorName = 1, ..., ColorVocabularyCardinality, can be community-agreed upon based on various kinds of evidence, whether viewed all at once or over time, such as a combination of prior beliefs with additional evidence inferred from new data in agreement with a Bayesian updating rule, largely applied in AI, including design and implementation of expert systems. In Bayesian updating, Bayesian inference is applied iteratively (Ghahramani, 2011; Wikipedia, 2018c): after observing some evidence, the resulting posterior probability can be treated as a prior probability and a new posterior probability computed from new evidence. A binary relationship R: \( A \rightarrow B \subseteq A \times B \), where sets A and B are categorical variables generated from a single population, guides the interpretation process of a two-way contingency table, also known as association matrix, cross tabulation, bivariate table or bivariate frequency table (BIVRFTAB) (Kuzera & Pontius, 2008; Pontius & Connors, 2006; Pontius & Millones, 2011), such that \( \text{BIVRFTAB} = \text{FrequencyCount}(A \times B) \). In the conventional domain of frequentist inference with no reference to prior beliefs, a BIVRFTAB is the 2-fold Cartesian product \( A \times B \) instantiated by the bivariate frequency counts of the two univariate categorical variables A and B generated from a single population. Hence, binary relationship R: \( A \rightarrow B \subseteq A \times B \neq \).
FrequencyCount(A × B) = BIVRFTAB, where one instantiation of the former guides the interpretation process of the latter. In greater detail, for any BIVRFTAB instance, either square or non-square, there is a binary relationship R: A ⇒ B ⊆ A × B that guides the interpretation process, where “correct” binary entry-pair cells of the 2-fold Cartesian product A × B are equal to 1 and located either off-diagonal (scattered) or on-diagonal, if a main diagonal exists when the BIVRFTAB is square. When a BIVRFTAB is estimated from a geospatial population with or without sampling, it is called overlapping area matrix (OAMTRX) (Baraldi et al., 2014; Baraldi, Bruzzone, & Blonda, 2005; Baraldi et al., 2006; Beauchemin & Thomson, 1997; Lunetta & Elvidge, 1999; Ortiz & Oliver, 2006; Pontius & Connors, 2006). When the binary relationship R: A ⇒ B is a bijective function (both 1–1 and onto), i.e., when the two categorical variables A and B estimated from a single population coincide, then the BIVRFTAB instantiation is square and sorted; it is typically called confusion matrix (CMTRX) or error matrix (Congalton & Green, 1999; Lunetta & Elvidge, 1999; Pontius & Millones, 2011; Stehman & Czaplewski, 1998). In a CMTRX, the main diagonal guides the interpretation process. For example, a square OAMTRX = FrequencyCount(A × B), where A = test thematic map legend, B = reference thematic map legend such that cardinality a = b, is a CMTRX if and only if A = B, i.e., if the test and reference codebooks are the same sorted set of concepts or categories. In general the class of (square and sorted) CMTRX instances is a special case of the class of OAMTRX instances, either square or non-square, i.e., OAMTRX ⊃ CMTRX. A similar consideration holds about summary Q^2Is generated from an OAMTRX or a CMTRX, i.e., Q^2I (OAMTRX) ⊃ Q^2I(CMTRX) (Baraldi et al., 2014, 2005, 2006).

Equation (3) shows that for any spatial unit x in the image-domain, when a hierarchical CV classification approach estimates posterior m(c| ColorValue(x), ShapeValue(x), TextureValue(x), SpatialRelationships(x, Neigh(x))) starting from an a priori knowledge-based near real-time color naming first stage, where condition m(ColorValue(x)| ColorName*) = 1 holds, if condition m(ColorName*| c) = 0 is true according to a static community-agreed binary relationship R: VocabularyOfColorNames ⇒ LegendOfObjectClassNames (and vice versa) known a priori, see
Table 3, then m(c| ColorValue(x), ShapeValue(x), TextureValue(x), SpatialRelationships(x, Neigh(x))) = 0 irrespective of any second-stage assessment of spatial terms ShapeValue(x), TextureValue(x) and SpatialRelationships(x, Neigh(x)), whose computational model is typically difficult to find and computationally expensive. Intuitively, Equation (3) shows that static color naming of any spatial unit x, either (0D) pixel, (1D) line or (2D) polygon, allows the color-based stratification of unconditional multivariate spatial variables into color class-conditional data distributions, in agreement with the statistic stratification principle (Hunt & Tyrrell, 2012) and the divide-and-conquer (dividi-et-impera) problem solving approach (Bishop, 1995; Cherkassky & Mulier, 1998; Lipson, 2007). Well known in statistics, the principle of statistic stratification guarantees that “stratification will always achieve greater precision provided that the strata have been chosen so that members of the same stratum are as similar as possible in respect of the characteristic of interest” (Hunt & Tyrrell, 2012).

Whereas 3D color polyhedra are easy to visualize and intuitive to think of in a true- or false-color RGB data cube, see Figure 7, hyperpolyhedra are difficult to think of and impossible to visualize in a MS reflectance space whose spectral dimensionality MS >3, with spectral channels ranging from visible to thermal portions of the electromagnetic spectrum, see Figure 1. Since it is non-adaptive-to-data, any static hyperpolyhedralization of a MS measurement space must be based on a priori physical knowledge available in addition to sensory data. Equivalent to a physical data model, static hyperpolyhedralization of a MS data space requires all spectral channels to be provided with a physical unit of radiometric measure, i.e., MS data must be radiometrically calibrated, in compliance with the GEO-CEOS QA4EO Cal requirements (GEO-CEOS, 2010, 2010), refer to Chapter 1.

Noteworthy, sensory data provided with a physical unit of measure can be input to both statistical/inductive and physical/deductive models, including hybrid (combined deductive and inductive) inference systems, refer to Chapter 1. On the contrary, uncalibrated dimensionless sensory data can be input to statistical data models exclusively. Although considered mandatory by the GEO-CEOS QA4EO Cal guidelines (GEO-CEOS, 2010) and regarded as a well-known “prerequisite for physical model-based analysis of airborne and satellite sensor measurements in the optical domain” (Schaepman-Strub, Schaepman, Painter, Dangel, & Martonchik, 2006), EO data Cal is ignored by relevant portions of the RS literature focusing on statistical EO data analytics, such as supervised learning-from-data function regression and classification (Bishop, 1995; Cherkassky & Mulier, 1998). One consequence is that, to date, statistical model-based EO-IUSs dominate the RS literature as well as commercial EO image processing software toolboxes, which typically consist of overly complicated collections of inherently ill-posed inductive machine learning-from-data algorithms (Bishop, 1995; Cherkassky & Mulier, 1998) to choose from based on heuristics (Baraldi, 2017; Baraldi & Boschetti, 2012a, 2012b). This is in contrast with the hybrid inference framework adopted by all biological cognitive systems (Parisi, 1991). In the words of O. Etzioni, “with all due respect to (machine learning-from-data scientists), thought is not a vector, and AI is not a problem in statistics’ (Etzioni, 2017; Marcus, 2018).

3. Related works in static MS reflectance space hyperpolyhedralization
In the RS discipline, there is a long history of hybrid EO-IUSs in operating mode, suitable for either low-level EO image enhancement (pre-processing) or high-level EO image understanding (classification), where an a priori knowledge-based decision tree for static MS reflectance space hyperpolyhedralization is plugged into the hybrid CV system architecture without Val by independent means, in disagreement with the GEO-CEOS QA4EO Val requirements (GEO-CEOS, 2010; Group on Earth Observation/Committee on Earth Observation Satellites (GEO-CEOS WGCV, 2015), refer to Chapter 1.

In recent years, the SIAM stratification of single-date MS imagery into MS color names was applied to MS image topographic correction, which is a traditional chicken-and-egg dilemma (Bishop & Colby, 2002; Bishop et al., 2003; Riano et al., 2003), synonym of inherently ill-posed problem in the Hadamard sense (Hadamard, 1902). When an inherently ill-posed MS image topographic correction was better conditioned for numeric solution by a prior knowledge-based SIAM color naming (masking) first stage, it required no human-machine interaction to run (Baraldi...
et al., 2010c), in compliance with process requirements of systematic ESA EO Level 2 product generation (ESA, 2015; DLR & VEGA, 2011; CNES, 2015), see Figure 12.

In the Atmospheric/Topographic Correction for Satellite Imagery (ATCOR) commercial software product, several deductive spectral pixel-based decision trees are implemented for use in different stages of an EO data enhancement pipeline (Baraldi et al., 2013; Baraldi & Humber, 2015; Dorigo et al., 2009; Richter & Schläpfer, 2012a, 2012b; Schläpfer, Richter & Hueni, 2009), see Figure 11. One of the ATCOR’s prior knowledge-based per-pixel decision trees delivers as output a haze/cloud/water (and snow) classification mask file (“image_out_hcw.bsq”). In addition, ATCOR includes a so-called prior knowledge-based decision tree for Spectral Classification of surface reflectance signatures (SPECL) (Baraldi & Humber, 2015; Baraldi et al., 2013; Dorigo et al., 2009), see Table 4. Unfortunately, SPECL has never been tested by its authors in the RS literature, although it has been validated by independent means (Baraldi & Humber, 2015; Baraldi et al., 2013).

Supported by NASA, atmospheric effect removal by the Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS) project relies on exclusion masks for water, cloud, shadow and snow surface types detected by a simple set of prior knowledge-based spectral decision rules applied per pixel. Quantitative analyses of LEDAPS products led by its authors revealed that these exclusion masks are prone to errors, to be corrected in future LEDAPS releases (Vermote & Saleous, 2007). Unfortunately, to date, in a recent comparison of cloud and cloud-shadow detectors, those implemented in LEDAPS scored low among alternative solutions (Foga et al., 2017).

In the 1980s, to provide an automatic alternative to a visual and subjective assessment of the cloud cover on Advanced Very High Resolution Radiometer (AVHRR) quicklook images in the ESA Earthnet archive, Muirhead and Malkawi developed a simple algorithm to classify daylight AVHRR images on a pixel-by-pixel basis into land, cloud, sea, snow or ice and sunglint, such that the classified quicklook image was presented in appropriate pseudo-colors, e.g., green: land, blue: sea, white: cloud, etc. (Muirhead & Malkawi, 1989).

Developed independently by NASA (Ackerman et al. 1998) and the Canadian Center for Remote Sensing (CCRS) (Luo et al., 2008), pixel-based static decision trees contribute, to date, to the systematic generation of clear-sky Moderate Resolution Imaging Spectroradiometer (MODIS) image composites in operating mode, see Figure 17.

To pursue high-level LC/LCC detection through time, extensions to the time domain of a single-date a priori spectral rule base for VQ of an MS reflectance space have become available to the general public in 2015 through the Google Earth Engine (GEE) platform (Simonetti et al. 2015b) or in the form of a commercial LC/LCC map product at national scale (GeoTerraImage, 2015). These are both post-classification approaches for LC/LCC detection based on a time-series of single-date per-pixel prior knowledge-based MS decision-tree classification maps. In practice, MS image time-series analysis in the domain of numeric sub-symbolic variables is replaced by MS color map time-series analysis in the domain of categorical semi-symbolic variables. These two post-classification approaches share the same operational limitations, specifically, they are Landsat sensor series-specific, pixel-based, where spatial topological and spatial non-topological information components are totally ignored, and their post-classification overall accuracy (OA) ∈ [0, 1] is not superior to the product of the single-date classification accuracies in the time-series (Lunetta & Elvidge, 1999). For example,

\[
\text{Bi-temporal post-classification LC change/no-change (LCC) detection overall accuracy (OA)} = \frac{\text{OA-LCC}_{1} \times \text{OA-LCC}_{2}}{2} \leq \text{OA-LCC}_{1} \times \text{OA-LCC}_{2} \quad \text{where OA-LCC}_{1} \in [0, 1] \text{and OA-LCC}_{2} \in [0, 1]
\]

(4)
Table 4. Rule set (structural knowledge) and order of presentation of the rule set (procedural knowledge) adopted by the prior knowledge-based MS reflectance space quantizer called Spectral Classification of surface reflectance signatures (SPECL), implemented within the ATCOR commercial software toolbox (Dorigo et al., 2009; Richter & Schläpfer, 2012a, 2012b).

| 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, 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 |
| 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 AND 0.05 ≤ b5 ≤ 0.25 AND b5/b4 ≥ 0.7 |
| 11    | Sand/bare soil/cloud | b4/b3 ≤ 2.0 AND b4 ≥ 0.15 AND b5 ≥ 0.15 |
| 12    | Bright sand/bare soil/cloud | b4/b3 ≤ 2.0 AND b4 ≥ 0.15 AND (b4 ≥ 0.25 OR b5 ≥ 0.30) |              |
| 13    | Dry vegetation/soil | (1.7 ≤ b4/b3 ≤ 2.0 AND b4 ≥ 0.25 OR (1.4 ≤ b4/ b3 ≤ 2.0 AND b7/b5 ≤ 0.83) |              |
| 14    | Sparse veg./soil   | (1.4 ≤ b4/b3 ≤ 1.7 AND b4 ≥ 0.25 OR (1.4 ≤ b4/ b3 ≤ 2.0 AND b7/b5 ≤ 0.83 AND b5/b4 ≥ 1.2) |              |
| 15    | Turbid water        | b4 ≤ 0.11 AND b5 ≤ 0.05 |              |
| 16    | Clear water         | b4 ≤ 0.02 AND b5 ≤ 0.02 |              |
| 17    | Clear water over sand | b3 ≥ 0.02 AND b3 ≥ b4 + 0.005 AND b5 ≤ 0.02 |              |
| 18    | Shadow              |                               |              |
| 19    | Not classified (outliers) |                               |              |

These expressions are optional and only used if b5 is present. Decision rule depends on presence of b5. Decision rule depends on presence of b7.
In Equation (4), if OA-LC_1 = 0.90 and OA-LC_2 = 0.90, then OA-LCC_1,2 ≤ 0.81. Hence, post-classification analysis is recommended for its simplicity if and only if single-date OA values are “high” through the time-series (Baraldi, 2017; Baraldi et al., 2016; Tiede et al., 2016). In other words, a necessary not sufficient pre-condition for multi-temporal image analysis to score “high” in accuracy according to a conceptually simple and computationally efficient post-classification LC change/no-change detection approach is that single-date image classification accuracies score individually “high” through the time-series. The two aforementioned post-classification approaches were both inspired by a year 2006 SIAM instantiation of a static decision tree for Landsat reflectance space hyperpolyhedrization, presented in pseudo-code in the RS literature (Baraldi et al., 2006) and further developed into the SIAM application software available to date (Baraldi, 2017, 2011; Baraldi & Boschetti, 2012a, 2012b; Baraldi et al., 2014, 2010a, 2010b; Baraldi & Humber, 2015; Baraldi et al., 2013; Baraldi, Wassenaar, & Kay, 2010d).

In Boschetti et al. (2015), a year 2013 SIAM instantiation was successfully employed to accomplish post-classification burned area detection in MS image time-series.

Among the aforementioned static decision trees for MS color naming, only SIAM claims scalability to several families of EO imaging sensors featuring different spectral resolutions, see Table 1.

It is obvious but not trivial to emphasize to the RS community that, in human vision and CV, an a priori vocabulary of general-purpose data- and application-independent BC names is equivalent to a static sub-symbolic or semi-symbolic categorical variable non-coincident with...
a symbolic categorical variable whose levels are user- and application-specific classes of objects observed in the 4D spatio-temporal scene-domain, refer to Table 3 and Equation (3). The very same consideration holds for any discrete and finite set of spectral endmembers in mixed pixel analysis, which “cannot always be inverted to unique LC class names” (Adams et al., 1995). It means that spectral endmembers in hyperspectral (HS) image analysis are conceptually equivalent to a static, user- and application-independent vocabulary of BC names, corresponding to a mutually exclusive and totally exhaustive set of neither necessarily convex nor connected hyperpolyhedra in a hyperspectral color space. Whereas the SIAM expert system has been successfully applied to HS imagery for fully automated color naming and superpixel detection (Baraldi, 2017), in the RS literature spectral endmembers detection in HS imagery is traditionally dealt with by inductive learning-from-data algorithms (Ghamisi et al., 2017), which are typically site specific and semiautomatic (Liang, 2004).

Quite surprisingly, the non-coincident assumption between an a priori vocabulary of sub-symbolic color names A in the (2D) image-domain and an application-specific legend of symbolic classes B of real-world objects in the 4D scene-domain, where A ≠ B always holds true, refer to Table 3 and Equation (3), appears somehow difficult to acknowledge by relevant portions of the RS community. For example, in the DigitalGlobe Geospatial Big Data platform (GBDX), a patented prior knowledge-based decision tree for pixel-based very high resolution WorldView-2 and WorldView-3 image mapping onto static MS reflectance space hyperpolyhedra (GBDX Registered Name: protogenV2LULC, Provider: GBDX) was proposed to RS end-users as an “Automated Land Cover Classification” (DigitalGlobe, 2016). This program name can be considered somehow misleading because it refers to no EO image mapper of DNs into LC class names, but to a static sub-symbolic color space partitioner, where DNs are mapped onto color name-specific hyperpolyhedra, see Equation (3). Due to the confusion between color names in the (2D) image-domain with target LC classes in the 4D scene-domain, see Table 3, the “Automated Land Cover Classification” computer program is affected by several “known issues” (DigitalGlobe, 2016): “Vegetation: Thin cloud (cloud edges) might be misinterpreted as vegetation; Water: False positives maybe present due to certain types of concrete roofs or shadows; Soils: Ceramic roofing material and some types of asphalt may be misinterpreted as soil,” etc.

In Salmon et al. (2013), a year 2006 SIAM’s a priori dictionary of static sub-symbolic MS color names was downscaled in cardinality and sorted in the order of presentation to form a bijective function with a legend of symbolic classes of target objects in the scene-domain. In practice these authors forced a non-square BIVRFTAB to become a (square and sorted) CMTRX, where the main diagonal guides the interpretation process (Congalton & Green, 1999), to make it more intuitive and familiar to RS practitioners. In general, no binary relationship R: A ⇒ B between an a priori vocabulary A of static sub-symbolic color names and a user- and application-dependent dictionary B of symbolic classes of observed objects in the scene-domain is a bijective function, refer to Table 3 and Equation (3). As a consequence of its unrealistic hypothesis in color information/knowledge representation, the 1D image classification approach proposed in Salmon et al. (2013) scored low in accuracy. Unfortunately, to explain their poor MS image classification outcome these authors concluded that, in their experiments, a year 2006 SIAM’s static dictionary of color names was useless to identify target LC classes. The lesson to be gained by these authors’ experience is that well-established RS practices, such as 1D image analysis based on supervised data learning algorithms and thematic map quality assessment by means of a square and sorted CMTRX where test and reference thematic legends are the same, can become malpractices when an a priori dictionary of static color names is employed for MS image classification purposes in agreement with Equation (3) and common sense, see Table 3. This lesson learned is supported by the fact that one of the same co-authors of paper (Salmon et al., 2013) reached opposite conclusions when a year 2013 SIAM application software, the same investigated by the present paper, was employed successfully in detecting burned areas from MS image time-series according to a convergence of color names with spatio-temporal visual properties in agreement with Equation (3) (Boschetti et al., 2015).
4. Original hybrid eight-step guideline for identification of a categorical variable-pair binary relationship

Summarized in Chapter 1, our experimental project required to compare an annual time-series of test SIAM-WELD maps of sub-symbolic color names, see Figure 6, with a reference USGS NLCD 2006 map whose legend of symbolic LC classes is summarized in Table 2. Since these test and reference map legends do not coincide, they must be reconciled/harmonized through a binary relationship \( R: \text{VocabularyOfColorNames} \Rightarrow \text{LegendOfObjectClassNames} \) (and vice versa), refer to Equation (3).

The harmonization of ontologies and the comparison of thematic maps with different legends are the subject of research of a minor body of literature, e.g., refer to works in ontology-driven geographic information systems (ODGIS) (Fonseca, Egenhofer, Agouris, & Camara, 2002; Guarino, 1995; Sowa, 2000). Ahlqvist writes that “to negotiate and compare information stemming from different classification systems (Bishr, 1998; Mizen, Dolbear, & Hart, 2005)... a translation can be achieved by matching the concepts in one system with concepts in another, either directly or through an intermediate classification (Feng & Flewelling, 2004; Kavouras & Kokla, 2002)” (Ahlqvist, 2005). Stehman describes four common types of thematic map-pair comparisons (Stehman, 1999). In the first type, different thematic maps, either crisp or fuzzy, of the same region of interest and employing the same sorted set (legend) of LC classes are compared (Kuzera & Pontius, 2008). In the second type, which includes the first type as a special case, thematic maps, either crisp or fuzzy, of the same region of interest, but featuring map legends that differ in their basic terms with regard to semantics and/or cardinality and/or order of presentation are compared. The third and fourth types of thematic maps comparison regard maps of different surface areas featuring, respectively, the same dictionary or different dictionaries of basic terms. Whereas a large portion of the RS community appears concerned with the aforementioned first type of map comparisons exclusively, the protocol proposed in (Baraldi et al., 2014) focuses on the second type, which includes the first type as a special case. In Couclelis (2010), the author observed that inter-dictionary concept matching (“conceptual matching”) (Ahlqvist, 2005) is an inherently equivocal information-as-data-interpretation process (Capurro & Hjørland, 2003), see Table 3. In common practice, two independent human domain-experts (cognitive agents, knowledge engineers) are likely to identify different binary associations between two codebooks of codewords (Laurini & Thompson, 1992). The conclusion is that no “universal best match” of two different codebooks can exist, but identification of the most appropriate binary relationship between two different nomenclatures becomes a subjective matter of negotiation to become community-agreed upon (Baraldi, 2017; Capurro & Hjørland, 2003; Couclelis, 2010).

To streamline the inherently subjective selection of “correct” entry-pairs in a binary relationship \( R: A \Rightarrow B \subseteq A \times B \) between two univariate categorical variables \( A \) and \( B \) estimated from a single population, an original hybrid eight-step guideline was designed for best practice, where deductive/top-down prior beliefs and inductive/bottom-up learning-from-data inference are combined. This hybrid protocol is sketched hereafter as the second original and pragmatic contribution of the present Part 1 to fill the gap from EO sensory data, mapped onto BC names, to ESA EO Level 2 product, whose SCM features semantics. As an example, let us consider a binary relationship \( R: A \Rightarrow B = \text{VocabularyOfColorNames} \Rightarrow \text{LegendOfObjectClassNames} \subseteq A \times B \) where rows are a test set of three semi-symbolic color names, say, \( A = \{\text{MS green-as-"Vegetation"}, \text{MS white-as-"Cloud"}, \text{"Unknowns"}\} \), where \( |A| = a = \text{ColorVocabularyCardinality} = TC = 3 \) is the row (test) cardinality, and where columns are a reference set of three symbolic LC classes, say, \( B = \{\text{"Evergreen Forest"}, \text{"Deciduous Forest"}, \text{"Others"}\} \), where \( |B| = b = \text{ObjectClassLegendCardinality} = RC = 3 \) is the column (reference) cardinality.

(1) Display multivariate frequency distributions of the two univariate categorical variables estimated from a single population in the BIVRFTAB = FrequencyCount(\( A \times B \)) whose size is \( TC \times RC \).
Switch any data-derived \( p(r \mid t) = 1 \), otherwise set that cell to 1. Let us identify this contingency table instantiation as \( \text{DataDrivenConditionalProb}(r \mid t)(x, y) = 1 \), \( \in \), \( TH1 \subseteq \) shows an example of how this protocol can be employed in practice. In Table TH1 or B the last activity occurs as late as possible in the information-as-data-interpretation process, where a test spectral category MS white-as-“Cloud” can match a reference LC class “Evergreen Forest”: this data-derived entry-pair match must be considered non-correct in the final R: A \( \times \) B following semantic scrutiny by a human expert.

Table 5 shows an example of how this protocol can be employed in practice. In Table 5 the last step 8 identifies an inherently equivocal information-as-data-interpretation process, where a human decision maker has a pro-active role in providing frequency data with semantics (symbolic meanings) (Capurro & Hjørland, 2003). It is highly recommended that any inherently subjective information-as-data-interpretation activity occurs as late as possible in the information processing workflow, to avoid propagation of “errors” due to personal preferences not yet community-agreed upon. Noteworthy, in the proposed eight-step guideline there are two “hidden” system’s free-parameters to be user-defined based on heuristics, equivalent to a trial-and-error strategy: variables TH1 and TH2 are two numeric thresholds in range \( [0, 1] \) for binary (hard, crisp) decision making, whose normalized range of change and intuitive meaning in terms of probability should make their selection easy and, to a certain extent, application- and user-independent.

5. Original measure of association (harmonization) in a categorical variable-pair binary relationship eligible for guiding the interpretation process of a two-way contingency table

Traditional scalar indicators of bivariate categorical variable association estimated from a BIVRFTAB = FrequencyCount(A \( \times \) B), either square or non-square, include the Pearson’s chi-square index of statistical independence and the normalized Pearson’s chi-square index, also known as Cramer’s coefficient V (Sheskin, 2000). These frequentist statistics of independence do not apply to a binary relationship R: A \( \Rightarrow \) B \( \subseteq \) A \( \times \) B, such as that shown in Table 3, where there is no frequency count, i.e., binary relationship R: A \( \Rightarrow \) B \( \subseteq \) A \( \times \) B \( \neq \) FrequencyCount(A \( \times \) B) = BIVRFTAB, refer to Chapter 4. Hereafter, a scalar indicator of association (harmonization, reconciliation) between two
Table 5. 8-step guideline for best practice in the identification of a dictionary-pair binary relationship based on a hybrid combination of (top-down) prior beliefs, if any, with (bottom-up) frequentist inference

**STEP 1. Dictionary-pair relationship, multivariate occurrence distributions**

| Test Classification (TC) | Reference Classification (RC) |
|--------------------------|-------------------------------|
|                          | EvergreenF | DeciduousF | Others |
| **Vegetation**           | 10         | 30         | 60     | 100    |
| **Cloud**                | 2          | 0          | 10     | 12     |
| **Unknowns**             | 0          | 5          | 100    | 105    |
| **Tot.**                 | 12         | 35         | 170    | 217    |

**STEP 2. Dictionary-pair relationship, multivariate probability distributions**

| Test Classification (TC) | Reference Classification (RC) |
|--------------------------|-------------------------------|
|                          | EvergreenF | DeciduousF | Others |
| **Vegetation**           | 0.046082949 | 0.138248848 | 0.276498 |
|                          | 0.460829    |            |        |
| **Cloud**                | 0.00921659  | 0          | 0.046083 |
|                          | 0.0553      |            |        |
| **Unknowns**             | 0           | 0.023041475 | 0.460829 |
|                          | 0.483871    |            |        |
| **Tot.**                 | 0.055299539 | 0.161290323 | 0.78341 |

**STEP 3. Dictionary-pair relationship, cond. prob. (RC|TC)**

| Test Classification (TC) | Reference Classification (RC) |
|--------------------------|-------------------------------|
|                          | EvergreenF | DeciduousF | Others |
| **Vegetation**           | 0.1         | 0.3        | 0.6    | 1     |
| **Cloud**                | 0.166666667 | 0          | 0.833333 | 1     |
| **Unknowns**             | 0           | 0.047619048 | 0.952381 |
|                         |             |            |        | 1     |

**STEP 4. Crisp membership function (RC|TC) > TH1 = 0.09.**

| Test Classification (TC) | Reference Classification (RC) |
|--------------------------|-------------------------------|
|                          | EvergreenF | DeciduousF | Others |
| **Vegetation**           | 1           | 1          | 1     |
| **Cloud**                | 1           | 0          | 1     |
| **Unknowns**             | 0           | 0          | 1     |

**STEP 5. Dictionary-pair relationship, cond. prob. (TC|RC)**

| Test Classification (TC) | Reference Classification (RC) |
|--------------------------|-------------------------------|
|                          | EvergreenF | DeciduousF | Others |
| **Vegetation**           | 0.833333333 | 0.857142857 | 0.352941 |
| **Cloud**                | 0.166666667 | 0          | 0.058824 |
| **Unknowns**             | 0           | 0.142857143 | 0.588235 |
|                         | 1           | 1          | 1     |

(Continued)
univariate categorical variables (codebooks of codewords), A and B, collected from a single population, called Categorical Variable Pair Association Index (CVPAI) in range $[0, 1]$, is estimated from a binary relationship $R: A \Rightarrow B$, such that $CVPAI(R: A \Rightarrow B) \in [0, 1]$.

Proposed in Baraldi et al. (2014), a CVPAI version 1, $CVPAI_1(R: A \Rightarrow B) \in [0, 1]$, is maximized (tends to 1, meaning maximum harmonization) if the binary relationship $R: A \Rightarrow B$ from set $A =$ test categorical variable, e.g., vocabulary of color names, to set $B =$ reference categorical variable, e.g., dictionary of LC class names, is a bijective function, i.e., the binary relationship $R: A \Rightarrow B$ is a function, therefore to each instance in test set $A$ (of color names) corresponds a single instance in reference set $B$ (of LC class names), and this function is both injective (one-to-one, for any instance in reference set $B$ of LC class names there is no more than one instance in test set $A$ of color names) and surjective (onto, for any instance in reference set $B$ of LC class names there is at least one instance in test set $A$ of color names), see Figure 18.

Hereafter, original formulations of CVPAI version 2, $CVPAI_2(R: A \Rightarrow B) \in [0, 1]$, and CVPAI version 3, $CVPAI_3(R: A \Rightarrow B) \in [0, 1]$, complementary not alternative to the CVPAI1 formulation presented in (Baraldi et al., 2014), are proposed as the third original and analytic contribution of the present Part 1 of this paper. Unlike the CVPAI1 expression, a novel CVPAI2 formulation was constrained as follows, see Figure 18. (i) The “most discriminative” test-to-reference inter-set binary relation $R: A \Rightarrow B$ is a function, i.e., each test color name in test set $A$ matches with only one reference LC class name in reference set $B$. (ii) The “most discriminative” reference-to-test class relation is either a surjective function, i.e., each reference LC class in set $B$ matches with at least one test color name in set $B$, or a bijective function, both surjective and injective as a special case of the former, i.e., each reference LC class in set $B$ matches with only one test color name in set $A$, see Figure 18. In short, $CVPAI_2(R: A \Rightarrow B)$ is

Table 5. (Continued)

| STEP 6. Crisp membership function (TC|RC) > TH2 = 0.06 ≤ TH1 = 0.09. | Reference Classification (RC) |
|--------------------------------------------------------|--------------------------------|
| Test Classification (TC) | EvergreenF | DeciduousF | Others |
| Vegetation               | 1          | 1          | 1      |
| Cloud                    | 1          | 0          | 0      |
| Unknowns                 | 0          | 1          | 1      |

| STEP 7. OR(Crisp membership function (TC|RC), Crisp membership function(RC|TC)) | Reference Classification (RC) |
|--------------------------------------------------------|--------------------------------|
| Test Classification (TC) | EvergreenF | DeciduousF | Others |
| Vegetation               | 1          | 1          | 1      |
| Cloud                    | 1          | 0          | 1      |
| Unknowns                 | 0          | 1          | 1      |

| STEP 8. Top-down (driven-by-prior knowledge) scrutiny of bottom-up (data-driven) “temporary correct” or “temporary non-correct” cells | Reference Classification (RC) |
|--------------------------------------------------------|--------------------------------|
| Test Classification (TC) | EvergreenF | DeciduousF | Others |
| Vegetation               | 1          | 1          | 1      |
| Cloud                    | 0          | 0          | 1      |
| Unknowns                 | 0          | 0          | 1      |
maximum (tends to 1, meaning maximum harmonization) when the binary relationship \( R: A \Rightarrow B \) is either a surjective function or a bijective function. Since the CVPAI2 formulation relaxes the CVPAI1 formulation, see Figure 18, it is always true that CVPAI2(\( R: A \Rightarrow B \)) ≥ CVPAI1(\( R: A \Rightarrow B \)) ∈ [0, 1]. The two formulations CVPAI1 and CVPAI2 are complementary not alternative because they have been designed to be maximized by different distributions of “correct” entry-pair cells in a binary relationship \( R: A \Rightarrow B \subseteq A \times B \), see Figure 18. Both formulations CVPAI1(\( R: A \Rightarrow B \)) and CVPAI2(\( R: A \Rightarrow B \)) are independent of frequency counts generated by a bivariate categorical variable distribution to be displayed in a BIVRFTAB = FrequencyCount(\( A \times B \)) ≠ R: \( A \Rightarrow B \subseteq A \times B \). Depending on the problem at hand, where two specific instantiations of test set A and reference set B are estimated from the same population, either CVPAI1 or CVPAI2 can be considered better suited for estimating the degree of harmonization between sets A and B, see Figure 18. If test set A is a vocabulary of BC names and reference test B is a legend of LC class names, then the novel CVPAI2 formulation is recommended for use. An alternative formulation to CVPAI2(\( R: A \Rightarrow B \)) is CVPAI3(\( R: A \Rightarrow B \)). They are both maximized by the same distribution of “correct” entry-pair cells in a binary relationship \( R: A \Rightarrow B \subseteq A \times B \).

The analytic formulation of CVPAI2(\( R: A \Rightarrow B \)), see Figure 18, is proposed as follows. In a binary relationship \( R: A \Rightarrow B \subseteq A \times B \), set A is a test codebook of cardinality \(|A| = TC\) as rows and set B is a reference codebook of cardinality \(|B| = RC\) as columns, so that the size of the 2-fold Cartesian product \( A \times B \) is \( TC \times RC \). The total number of “correct” entry-pair cells in \( R: A \Rightarrow B \) is identified as \( CE \), where \( 0 \leq CE \leq TC \times RC \). In addition, symbol “\( = \)" is adopted to mean “equal to". The CVPAI2 formulation is constrained as follows.

\[
(a) \quad CE = \sum_{t=1}^{TC} \sum_{r=1}^{RC} CE_{t,r}, \text{ with } CE_{t,r} \in \{0,1\} = \\
\{\text{non-correct entry} - \text{pair}(t,r)=0, \text{correct entry} - \text{pair}(t,r)=1\}, \quad CE \in \{0, RC \times TC\}.
\]

(b) If \( CE = 0 \) then CVPAI2 = 0 must hold. It means that, when no “correct” cell exists, then the degree of conceptual match between the two categorical variables is zero.

(c) If \( CE = TC \times RC \) then CVPAI2 ≠ 0 must hold. It means that when all table cells are considered “correct”, then no entry-pair is discriminative (informative), i.e., nothing makes the difference between the two categorical variables.

(d) If
\[
\left\{ \left[ \sum_{t=1}^{TC} CE_{t,r} = CE_{t,0} \right] > 0, r = 1, \ldots, RC \right\} \text{ AND } \\
\left\{ \sum_{r=1}^{RC} CE_{t,r} = CE_{t,0} \right\} = 1, t = 1, \ldots, TC,
\]

Figure 18. Entity-relationship conceptual model representation of the binary relationship \( R: A \Rightarrow B \) from set \( A = \text{test categorical variable to set } B = \text{reference categorical variable}, \text{provided with the min: max cardinality required by the Categorical Variable-Pair Association Index (CVPAI) formulation 1 (CVPAI1) and formulation 2 (CVPAI2) to score maximum in range [0, 1]. Inequality } CVPAI1 \leq CVPAI2 \text{ holds, i.e., the latter is a relaxed version of the former. In particular, CVPAI1 is maximum (equal 1) when the binary relationship } R: A \Rightarrow B \text{ from set } A = \text{test categorical variable to set } B = \text{reference categorical variable is a bijective function, both injective (one-to-one) and surjective (onto). CVPAI2 is maximum when the binary relationship } R: A \Rightarrow B \text{ is either a surjective function or a bijective function.} \]
where \( CE_{r,t} \) is the total sum of correct entry-cell pairs along reference column \( r \), with \( r = 1, \ldots, RC \), \( CE_{t} \) is the total sum of correct entry-cell pairs along test row \( t \), with \( t = 1, \ldots, TC \), then CVPAI2 must be maximum, i.e., \( CVPAI2 = 1 \). It means that if for each test class \( t = 1, \ldots, TC \) there is one single match and for each reference class \( r = 1, \ldots, RC \) there is at least one match, then CVPAI2 must be maximum, such that \( CVPAI2 = 1 \), see Figure 18.

(e) If [not condition(b) AND not condition(c) AND not condition(d)] then CVPAI2 \( \epsilon (0,1) \).

In a square binary relationship \( R: A \Rightarrow B \) where \( TC = RC \), to maximize the CVPAI2 (to become equal to 1), submitted to condition (d), the binary relationship must be a 1–1 function (an injective function, 1–1 forward and 1–1 backward). To satisfy the set of aforementioned constraints (a) to (e), the following set of original equations is proposed.

\[
CVPAI2 \in [0,1], \quad CVPAI2 = \frac{1}{RC + TC} \left( \sum_{r=1}^{RC} f_{RC}(CE_{+,r}) + \sum_{t=1}^{TC} f_{TC}(CE_{t,+}) \right). \tag{5}
\]

with

\[
f_{RC}(i) = \begin{cases} 0 & \text{if } i = 0, \\ 1 & \text{if } i > 0, \\ \end{cases} \quad i \in \{0, TC\} \subset I_{0}^{+}, \quad \text{where } i = CE_{+,r}, \quad r \in \{1, RC\}, \tag{6}
\]

\[
f_{TC}(j) = \begin{cases} 0 & \text{if } j = 0, \\ \text{GaussianMembership}(j, \text{Center} = 1, \text{StnDev} = RC/3) = \\ \quad e^{-\frac{j-1}{(\frac{RC}{3})^{2}}} \in [0,1], \quad \text{if } j > 0, \quad \text{with } j \in \{0, RC\} \subset I_{0}^{+}, \quad \text{where } j = CE_{+,t}, \quad t \in \{1, TC\}. \tag{7}
\]

Although it is maximized by the same distribution of “correct” entry-pair cells in a binary relationship \( R: A \Rightarrow B \subseteq A \times B \), a novel CVPAI3 expression is a more severe formulation of CVPAI2, i.e., \( 1 \geq CVPAI2 \geq CVPAI3 \in [0,1] \). The proposed CVPAI3 formulation alternative to CVPAI2 is the following.

\[
CVPAI3 \in [0,1], \quad CVPAI3 = \min \left\{ \frac{\sum_{r=1}^{RC} f_{RC}(CE_{+,r})}{RC}, \frac{\sum_{t=1}^{TC} f_{TC}(CE_{t,+})}{TC} \right\} \tag{8}
\]

In Equation (7), GaussianMembership\((j, \text{Center} = 1, \text{StnDev} = RC/3) \rightarrow 0 \) when it covers approximately 99.73% of the area underneath the Gaussian curve at distance \( j = \pm 3 \cdot \text{StnDev} = 3 \cdot RC/3 = \pm RC \) from its Center = 1. If \( j = 1 \), then GaussianMembership\((j, \text{Center} = 1, \text{StnDev} = RC/3) = 1 \). It is trivial to prove that Equations (5–7) satisfy the aforementioned requirements (a) to (d). By means of a numeric example, it can be shown that requirement (e) is satisfied too. For example, estimated from the binary relationship instantiated at step 8 of Table 5, CVPAI2 = \((1/6)^{*}(1 + 1 + 1 + 1 + 1 + \exp(-0.5*(3-1)^2/((3/3)^2))) = (1/6) \times (5 + 0.1353) = 0.8558 \). Intuitively closer to 1 than 0, this CVPAI2 value shows that the harmonization between the two test and reference nominal variables is (fuzzy) “high” (>0.8) in Table 5.

To appreciate the conceptual difference between the CVPAI1 and CVPAI2 formulations maximized by different distribution of “correct” entry-pairs in a binary relationship \( R: A \Rightarrow B \subseteq A \times B \), see Figure 18, let us compare a test vocabulary A of color names, such as SIAM’s, see Figure 6, with a reference dictionary B of LC class names, such as the USGS NLCD’s, see Table 2. In terms of capability of color names to discriminate LC class names, the ideal test-to-reference binary relationship is a function where one color name matches with only one reference LC class. On the other hand, the color attribute of a real-world LC class can be typically linked to one or more discrete color names, see Table 3. In this realistic example the range of change of an estimated CVPAI2 value would be \([0, 1]\), up to its maximum value
equal to 1, while the range of change of the CVPAI1 formulation proposed in (Baraldi et al., 2014) would be (0, 1), below its maximum value equal to 1, see Figure 18.

Another example where the difference between the CVPAI1 and CVPAI2 formulations is highlighted is when the test dictionary A is a specialized version of the reference dictionary B, according to a parent–child relationship. For example, a test taxonomy of LC classes is \(A = \text{LegendOfObjectClassNames}_A = \{\text{LC class } \text{"Dark-tone bare soil"}, \text{LC class } \text{"Light-tone bare soil"}, \text{LC class } \text{"Deciduous Forest"}, \text{LC class } \text{"Evergreen Forest"}\}\) and a reference LC class taxonomy is \(B = \text{LegendOfObjectClassNames}_B = \{\text{LC class } \text{"Dark-tone bare soil"}, \text{LC class } \text{"Bare soil"}, \text{LC class } \text{"Forest"}\}\). Based on our prior knowledge-based understanding of these two semantic dictionaries A and B, a reasonable binary relationship can be considered \(R: A \rightarrow B = \{(\text{LC class } \text{"Dark-tone bare soil"}, \text{LC class } \text{"Bare soil"}); (\text{LC class } \text{"Light-tone bare soil"}, \text{LC class } \text{"Bare soil"}); (\text{LC class } \text{"Deciduous Forest"}, \text{LC class } \text{"Forest"}); (\text{LC class } \text{"Evergreen Forest"}, \text{LC class } \text{"Forest"})\}\). In this case, the CVPAI1 formulation scores below its maximum, i.e., CVPAI1 \(\in (0, 1)\), while the expected CVPAI2 value would score maximum, i.e., CVPAI2 = 1, meaning that the two vocabularies are harmonized because one is the specialization of the other, featuring a parent–child relationship.

These two examples illustrate the intuitive meaning and practical use of the normalized quantitative indicator CVPAI2 \(\in [0, 1]\) in an EO-IUS implementation based on a convergence-of-evidence approach, in agreement with Equation (3). When the semantic information gap from sub-symbolic sensory data to a symbolic set \(B = \text{LegendOfObjectClassNames}\) is filled by an EO-IUS starting from a static color naming first stage, provided with a semi-symbolic set \(A = \text{VocabularyOfColorNames}\), if the binary relationship \(R: A \rightarrow B \subseteq A \times B\) features a degree of association CVPAI2 \(\in [0, 1]\), then \(1 – \text{Categorical CVPAI1} = 1\) is the semantic information gap from sub-symbolic sensory data to the symbolic LegendOfObjectClassNames left to be filled by further stages in the hierarchical EO-IUS pipeline, where spatial information is masked by first-stage color names. If CVPAI2 = 1, then secondary color information discretized by set \(A = \text{VocabularyOfColorNames}\) suffices to detect target set \(B = \text{LegendOfObjectClassNames}\) with no further need to investigate primary spatial information in a hierarchical convergence-of-evidence image classification approach, refer to Equation (3).

6. Conclusions
To pursue the GEO-CEOS visionary goal of a GEOSS implementation plan for years 2005–2015 not-yet accomplished by the RS community, this interdisciplinary work aimed at filling an analytic and pragmatic information gap from EO image big data to systematic ESA EO Level 2 product generation at the ground segment, never achieved to date by any EO data provider and postulated as necessary not sufficient pre-condition to GEOSS development. For the sake of readability this paper is split into two, the present Part 1 - Theory and the following Part 2 - Validation.

The original contribution of the present Part 1 is fourfold. A first lesson was learned from published works on prior knowledge-based MS reflectance space hyperpolyhedralization into static (non-adaptive-to-data) color names, according to the principle of color naming discovered by linguistics and investigated by CV in the realm of cognitive science, see Figure 13. In color naming, a static vocabulary of sub-symbolic color names is equivalent to a set of mutually exclusive and totally exhaustive (hyper)polyhedra, neither necessarily convex nor connected, in a color data (hyper)cube. It was observed that well-established RS practices, such as 1D image analysis based on supervised data learning algorithms, where dominant spatial information is neglected in favor of secondary color information, and thematic map quality assessment where test and reference map legends are required to coincide, can become mal-practices when an a priori dictionary of static color names is employed for MS image classification based on a convergence-of-evidence approach, such as in Bayesian naive classification, see Equation (3). When test and reference thematic map legends A and B are the same, the binary relationship \(R: A \rightarrow B \subseteq A \times B\) becomes a bijective function (both 1–1 and onto) and the main diagonal of the 2-fold Cartesian product \(A \times B\) guides the interpretation process of a bivariate frequency table, BIVRFTAB = FrequencyCount(A \(\times\) B), equal to a square and sorted confusion matrix, CMTRX. This constraint makes a CMTRX, whose input categorical variables A and B coincide, intuitive to understand and more familiar to RS practitioners. Noteworthy, inequality \(R: A \rightarrow B \not\subseteq A \times B\) = FrequencyCount(A \(\times\) B) = BIVRFTAB always holds true, where one
instance of the binary relationship $R$ guides the interpretation process of the two-way contingency table $\text{BIVRFTB}$. Quite surprisingly, the non-coincident assumption between an \textit{a priori} vocabulary $A$ of static sub-symbolic color names in the measurement color space and a user- and application-dependent legend $B$ of symbolic classes of real-world objects in the scene-domain, where inequality $A \neq B$ always holds, appears somehow difficult to acknowledge by relevant portions of the RS community, in contrast with common sense, see Table 3.

Second, Equation (3) was proposed as an analytic expression of a biologically plausible hybrid (combine deductive and inductive) CV system suitable for convergence of color and spatial evidence, in agreement with a Bayesian approach to vision proposed by Marr (1982), with the principle of statistic stratification well known in statistics and with the divide-and-conquer (divide-et-impera) problem solving criterion widely adopted in structured engineering. In compliance with common sense, see Table 3, Equation (3) shows that a static color naming first stage can be employed for stratification purposes of further spatial-context sensitive image classification stages. In the static color naming first stage, a binary relationship $R: A \Rightarrow B \subseteq A \times B$ from a vocabulary $A$ of general-purpose static color names in the MS color space to a taxonomy $B$ of LC class names in the 4D spatio-temporal scene-domain, such as the standard FAO LCCS taxonomy shown in Figure 3, can be established by human experts based on top-down prior beliefs, if any, in combination with bottom-up evidence inferred from new data, as described in Table 5. Once established and community-agreed upon, a binary relationship $R: A \Rightarrow B \subseteq A \times B$ from a vocabulary $A$ of static color names in a MS color space to a standard legend of LC classes in the 4D spatio-temporal scene-domain becomes equivalent to an \textit{a priori} knowledge base in a Bayesian updating framework, where Bayesian inference is applied iteratively: after observing some evidence, the resulting posterior probability can be treated as prior probability. For example, once community-agreed upon, Table 3 becomes equivalent to an \textit{a priori} knowledge base available in addition to sensory data, and a new posterior probability can be computed from new data, e.g., to pursue image classification in agreement with Equation (3).

Third, for best practice a hybrid eight-step protocol, sketched in Table 5, was proposed to infer a binary relationship, $R: A \Rightarrow B \subseteq A \times B$, from categorical variable $A$ to categorical variable $B$ estimated from the same population, where codebooks $A$ and $B$ can differ in cardinality, semantics or in the order of presentation of codewords. This protocol streamlines a hybrid combination of deductive prior beliefs by human domain experts with inductive evidence from data. It is of practical use because identification of a binary relationship $R: A \Rightarrow B$ is mandatory to guide the interpretation process of a bivariate frequency table, $\text{BIVRFTAB} = \text{FrequencyCount}(A \times B)$, where $A \neq B$ in general. Only if $A = B$ then $\text{BIVRFTAB}$ becomes equivalent to the well-known square and sorted CMTRX, where the main diagonal guides the interpretation process.

Fourth, in compliance with the GEO-CEOS QA4EO Val guidelines, two original and alternative formulations, $\text{CVPAI2}(R: A \Rightarrow B \subseteq A \times B) \in [0, 1]$ and $\text{CVPAI3}(R: A \Rightarrow B \subseteq A \times B) \in [0, 1]$, were proposed as categorical variable-pair degree of association (harmonization) in a binary relationship, $R: A \Rightarrow B$, from categorical variable $A$ to categorical variable $B$ estimated from the same population, where $A \neq B$ in general. When $\text{CVPAI2}$ or $\text{CVPAI3}$ is maximum, equal to 1, then the two categorical variables $A$ and $B$ are considered fully harmonized.

To comply with the GEO-CEOS QA4EO Cal/Val requirements, the subsequent Part 2 of this paper presents and discusses a GEO-CEOS stage 4 Val of the annual SIAM-WELD map time-series for the years 2006 to 2009 in comparison with the reference USGS NLCD 2006 map, based on an original protocol for wall-to-wall inter-map comparison without sampling where the test and reference maps feature the same spatial resolution and spatial extent, but whose legends are not the same and must be harmonized.

**Abbreviations**

| Abbreviation | Description |
|--------------|-------------|
| AI           | Artificial Intelligence |
| ATCOR        | Atmospheric/Topographic Correction commercial software product |
| AVHRR        | Advanced Very High Resolution Radiometer |
MIR: Medium InfraRed
MODIS: Moderate Resolution Imaging Spectroradiometer
MS: Multi-Spectral
NASA: National Aeronautics and Space Administration
NIR: Near InfraRed
NLCD: National Land Cover Data
NOAA: National Oceanic and Atmospheric Administration
OA: Overall Accuracy
OAMTRX: Overlapping Area Matrix
OBIA: Object-Based Image Analysis
ODGIS: Ontology-Driven GIS
OGC: Open Geospatial Consortium
OP: Outcome (product) and Process
OP-Q³I: Outcome and Process Quantitative Quality Index
QA: Quality Assurance
QA4EO: Quality Accuracy Framework for Earth Observation
Q²I: Quantitative Quality Indicator
R&D: Research and Development
RGB: monitor-typical Red-Green-Blue data cube
RMSE: Root Mean Square Error
RS: Remote Sensing
SCM: Scene Classification Map
SEN2COR: Sentinel 2 (atmospheric) Correction Prototype Processor
SIAM™: Satellite Image Automatic Mapper™
STRATCOR: Stratified Topographic Correction
SS: Super-Spectral
SURF: Surface Reflectance
TIR: Thermal InfraRed
TOA: Top-Of-Atmosphere
TOARF: TOA Reflectance
TM (superscript): (non-registered) Trademark
UML: Unified Modeling Language
USGS: US Geological Survey
Val: Validation
VQ: Vector Quantization
WELD: Web-Enabled Landsat Data set
WGCV: Working Group on Calibration and Validation

Acknowledgments
Prof. Ralph Maughan, Idaho State University, is kindly acknowledged for his contribution as active conservationist and for his willingness to share his invaluable photo archive with the scientific community as well as the general public. Andrea Baraldi thanks Prof. Raphael Capurro, Hochschule der Medien, Germany, and Prof. Christopher Justice, Chair of the Department of Geographical Sciences, University of Maryland, College Park, MD, for their support. Above all, the authors acknowledge the fundamental contribution of Prof. Luigi Boschetti, currently at the Department of Forest, Rangeland and Fire Sciences, University of Idaho, Moscow, Idaho, who conducted by independent means all experiments whose results are proposed in this validation paper. The authors also wish to
thank the Editor-in-Chief, Associate Editor and reviewers for their competence, patience and willingness to help.

### Funding

To accomplish this work Andrea Baraldi was supported in part by the National Aeronautics and Space Administration (NASA) under Grant No. NNX07AV19G issued through the Earth Science Division of the Science Mission Directorate. Andrea Baraldi, Dirk Tiede and Stefan Lang were supported in part by the Austrian Science Fund (FWF) through the Doctoral College GIScience (DK W1237-N23).

### Competing interests

As a source of potential competing interests of professional or financial nature, coauthor Andrea Baraldi reports he is the sole developer and intellectual property right (IPR) owner of the Satellite Image Automatic Mapper™ (non-registered trademark) computer program validated by an independent third-party in this research and licensed by the one-man-company Baraldi Consultancy in Remote Sensing (siam.andreabaraldi.com) to academia, public institutions and private companies, eventually free-of-charge. Throughout his scientific career Andrea Baraldi has put in place approved plans for managing any potential conflict arising from his IPR ownership of the SIAM™ computer software.

### Author details

Andrea Baraldi,1,2,3,4 E-mail: andreob311@gmail.com

ORCID ID: http://orcid.org/0000-0001-5196-9944

Michael Laurence Humber2

E-mail: mhumber@umd.edu

Dirk Tiede1

E-mail: dirk.tiede@sbg.ac.at

ORCID ID: http://orcid.org/0000-0002-5473-3344

Stefan Lang3

E-mail: stefan.lang@sbg.ac.at

1 Department of Agricultural Sciences, University of Naples Federico II, Portici, Italy.

2 Department of Geographical Sciences, University of Maryland, College Park, MD, USA.

3 Department of Geoinformatics – Z_GIS, University of Salzburg, Salzburg, Austria.

4 Italian Space Agency (ASI), Rome, Italy.

### Citation information

Cite this article as: GEO-CEOS stage 4 validation of the Satellite Image Automatic Mapper lightweight computer program for ESA Earth observation level 2 product generation - Part 1: Theory, Andrea Baraldi, Michael Laurence Humber, Dirk Tiede & Stefan Lang, Cogent Geoscience (2018), 4: 1467357.

### Cover image

Source: Satellite Image Automatic Mapper (SIAM) multi-level map generated automatically, without human-machine interaction, and in near real-time, specifically, in linear time complexity with image size, from the 30 m resolution annual Web-Enabled Landsat Data (WELD) image composite for the year 2006 of the conterminous U.S. (CONUS), radiometrically calibrated into top-of-atmosphere reflectance (TOARF) values. The multi-level map’s legend, shown at bottom left, is the SIAM intermediate discretization level, consisting of 48 basic color (BC) names reassembled into 19 spectral macro-categories by an independent human expert, according to a proposed hybrid (combined deductive and inductive) eight-step protocol for identification of a categorical variable-pair binary relationship, from a vocabulary of BC names to a dictionary of land cover (LC) class names. No discrete and finite vocabulary of color names, such as SIAM’s, equivalent to a set of mutually exclusive and totally exhaustive (hyper)polyhedra, neither necessarily convex nor connected, in a MS reflectance (hyper)space, should ever be confused with a symbolic (semantic) taxonomy of LC class names in the 4D geospace-time scene-domain. Black lines across the SIAM-WELD 2006 color map represent the boundaries of the 86 Environmental Protection Agency (EPA) Level III ecoregions of the CONUS, suitable for regional-scale statistical stratification required to intercept geospatial non-stationary statistics, typically lost when a global spatial average, e.g., at continental spatial extent, is superimposed on the local computational processes.

### Notes

Figure 2. Graphical representation of a dependence relationship part-of-denoted with symbol ‘⊂’-pointing from the supplier to the client in agreement with the standard Unified Modeling Language (UML) for graphical modeling of object-oriented software (Fowler, 2003), between computer vision (CV), whose special case is EO image understanding (EO-IU) in operating mode, where relationship subset-of, denoted with symbol “⊂” meaning specialization with inheritance from the superset to the subset, holds true, and a Global Earth Observation System of Systems (GEOSS) (GEO, 2005), such that “NASA EO Level 2 product → ESA EO Level 2 product ∈ EO-IU in operating mode ∈ CV → GEOSS” Synchrony of 4D spatiotemporal scene from (2D) image reconstruction and understanding, vision is acknowledged to be a cognitive problem very difficult to solve because: (I) non-polynomial (NP)-hard in computational complexity (Frintrop, 2011; Tsotsos, 1990), (ii) inherently ill-posed in the Hadamard sense, as it is affected by: (I) a 4D-to-2D data dimensionality reduction from the scene- to the image-domain, e.g., responsible of occlusion phenomena, and (II) a semantic information gap from ever-varying sub-symbolic sensory data (sensations) in the image-domain to stable symbolic percepts in the modeled world (mental world, world ontology, world model) (Fanseea et al., 2002; Laurini & Thompson, 1992; Matsuyama & Wang, 1990; Sonka et al., 1994; Sowa, 2003). A NASA Earth observation (EO) Level 2 product, defined as “a data-derived physical variable at the same resolution and location as Level 1 source data” (NASA 2016b), is part-of the ESA EO Level 2 product, defined as follows (ESA, 2015; DLR & VEGA, 2011): (a) a single-date multi-spectral (MS) image whose digital numbers (DNs) are radiometrically calibrated into surface reflectance (SURF) values corrected for atmospheric, adjacency and topographic effects, stacked with (b) its data-derived general-purpose, user- and application-independent scene classification map (SCM), whose thematic map legend includes quality layers cloud and cloud-shadow (CNES, 2015). Working hypothesis “NASA EO Level 2 product → ESA EO Level 2 ∈ EO-IU in operating mode → GEOSS” postulates that no GEOSS can exist if the necessary not sufficient pre-condition of systematic ESA EO Level 2 product generation is accomplished in advance as the mandatory first step in a hierarchical EO-IU workflow for scene-from-image reconstruction and understanding in operating mode. Figure 10. (a) Sentinel-2A (S2A) MSI Level-1C image radiometrically calibrated (Cal) into top-of-atmosphere reflectance (TOARF) values by the ESA data.
provider, depicting an Earth surface located south of the city of Salzburg, Austria. The city area is visible around the middle of the image upper boundary (Lat-long coordinates: 47°48.250'N 13°02.43.6'E). Acquired on 2015-09-11. Spatial resolution: 10 m. Image size: 110 × 110 km. TOARF values in range [0, 1] are byte-coded in range (0, 255). The MS image is shown in RGB false colors, where monitor channel R = Medium Infrared (MIR) = S2 Band 11, channel G = Near IR (NIR) = S2 Band 8, channel B = (visible) Blue = S2 Band 2. No histogram stretching is applied for visualization purposes. (b) L-SIAM color map at coarse color granularity, consisting of 18 spectral categories depicted in pseudo colors shown in the map legend. Coarse granularity color categories are generated by merging color hyperpolyhedra at fine color granularity, according to pre-defined parent-child relationships, refer to Table 1. (c) L-SIAM color map at fine color granularity, consisting of 96 spectral categories depicted in pseudo colors shown in the map legend, refer to Table 1. (d) Superpixels size and whose class-specific within-segment variance tends to be high. (e) Histogram stretching is applied for visualization purposes. (f) 8-adjacency cross-aura contour map in range {0, 255}. No histogram stretching is applied. (g) S9 6 color map at fine granularity. It shows contours of connected sets of pixels featuring the same color label. These connected-components are also called image-objects, segments or superpixels. (h) Per-pixel scalar difference between the input MS image shown in (a) and the superpixels. Mean = 2.60, Stdev = 3.45. Histogram stretching is applied for visualization purposes. (i) 8-adjacency cross-aura contour map in range {0, 255}. No histogram stretching is applied. (j) L-SIAM color map at fine color granularity, consisting of 96 spectral categories depicted in pseudo colors shown in the map legend, refer to Table 1. (k) L-SIAM color map at fine color granularity, consisting of 96 spectral categories depicted in pseudo colors shown in the map legend, refer to Table 1.

References

Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., & Susstrunk, S. (2011). SLIC superpixels compared to state-of-the-art superpixel methods. IEEE Transactions on Pattern Analysis Machine Intelligence, 1–8.

Ackerman, S. A., Strabala, K. I., Menzel, W. P., Frey, R. A., Moeller, C. C., & Gumley, L. E. (1998). Discriminating clear sky from clouds with MODIS. Journal of Geophysical Research, 103(S2), 141–157. doi:10.1029/1998JD200032

Adams, J. B., Donald, E. S., Kapos, V., Almeida Filho, R., Roberts, D. A., Smith, M. O., & Gillespie, A. R. (1995). Classification of multispectral images based on fractions of endmembers: Application to land-cover change in the Brazilian Amazon. Remote Sensing Environment, 52, 137–154. doi:10.1016/0034-4257(94)00098-8

Ahlvist, O. (2005). Using uncertain conceptual spaces to translate between land cover categories. International Journal of Geographical Information Science, 19, 831–857.

Ahlvist, O. (2008). In search of classification that supports the dynamics of science: The FAO Land Cover Classification System and proposed modifications. Environment and Planning B: Planning and Design, 35, 169–186. doi:10.1068/b3364

Arvor, D., Madiello, B. D., & Corpetti, T. (2016). Semantic pre-classification of vegetation gradient based on linearly unmixed Landsat time series. In Geoscience and Remote Sensing Symposium (IGARSS), IEEE International (pp. 4422–4425).

Baraldi, A. (2009). Impact of radiometric calibration and specifications of spaceborne optical imaging sensors on the development of operational automatic remote sensing image understanding systems. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2(2), 104–134. doi:10.1109/JSTARS.2009.2023801

Baraldi, A. (2011). Fuzzification of a crisp near-real-time operational automatic spectral-rule-based decision-tree preliminary classifier of multisource multi-spectral remotely sensed images. IEEE Transactions on Geoscience and Remote Sensing, 49, 2113–2134. doi:10.1109/TGRS.2010.2091137

Baraldi, A. (2015). “Automatic Spatial Context-Sensitive Cloud/Cloud-Shadow Detection in Multi-Source Multi-Spectral Earth Observation Images – AutoCloud+,” Invitation to tender ESA AO1-83731/15/NL – “VAE: Next Generation EO-based Information Services”, DOI: 10.13140/RG.2.2.34162.71363. arXiv: 1701.04256. Retrieved Jan., 8, 2017 from https://arxiv.org/ftp/arxiv/papers/1701/1701.04256.pdf

Baraldi, A., & Boschetti, L. (2012a). Operational automatic remote sensing image understanding systems: Beyond Geographic Object-Based and Object-Oriented Image Analysis (GEOBIA/GEOOIA) - Part 1: Introduction. Remote Sensing, 4, 2694–2735. doi:10.3390/rs4092694

Baraldi, A., & Boschetti, L. (2012b). 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 Sensing, 4, 2768–2817.

Baraldi, A., Boschetti, L., & Number, M. (2014). Probability sampling protocol for thematic and spatial quality assessment of classification maps generated from spaceborne/airborne very high resolution images. IEEE Transactions on Geoscience and Remote Sensing, 52(1), 701–760. doi:10.1109/TGRS.2013.2243739

Baraldi, A., Buzzzone, L., & Blond, P. (2005). Quality assessment of classification and cluster maps without ground truth knowledge. IEEE Transactions on Geoscience and Remote Sensing, 43, 857–873. doi:10.1109/TGRS.2004.843074
Baraldi, A., Durieux, L., Simonetti, D., Conchedda, G., Holecz, F., & Blonda, P. (2010a). Automatic spectral rule-based preliminary classification of radiometrically calibrated SPOT-4/-5/IRS, AVHRR/ MSG,AATSR, IKONOS/QuickBird/OrbView/GeoEye and DMSP/SPOT-1/-2 imagery - Part I: System design and implementation. IEEE Transactions on Geoscience and Remote Sensing, 48, 1299–1325. doi:10.1109/TGRS.2009.2032457

Baraldi, A., Durieux, L., Simonetti, D., Conchedda, G., Holecz, F., & Blonda, P. (2010b). “Automatic spectral rule-based preliminary classification of radiometrically calibrated SPOT-4/-5/IRS, AVHRR/ MSG,AATSR, IKONOS/QuickBird/OrbView/GeoEye and DMSP/SPOT-1/-2 imagery - Part II: Classification accuracy assessment,” IEEE Trans. IEEE Transactions on Geoscience and Remote Sensing, 48, 1326–1354. doi:10.1109/TGRS.2009.2032064.

Baraldi, A., Gironda, M., & Simonetti, D. (2010c). Operational two-stage stratified topographic correction of spaceborne multispectral imagery employing an automatic spectral rule-based decision-tree preliminary classifier. IEEE Transactions Geoscience and Remote Sensing, 48(1), 112–146. doi:10.1109/TGRS.2009.2028017

Baraldi, A. (2015). “Automatic Spatial Context-Sensitive Cloud/Cloud-Shadow Detection in Multi-Source Multi-Spectral Earth Observation Images – AutoCloud”, Invitation to tender ESA/AD/1-8373/15/I-NB – “VAE: Next Generation EO-based Information Services”. doi:10.13140/RG.2.2.34162.71363. arXiv:1701.04256. Retrieved Jan., 8, 2017 from https://arxiv.org/ftp/arxiv/papers/1701/1701.04256.pdf

Baraldi, A. (2015). “Automatic Spatial Context-Sensitive Cloud/Cloud-Shadow Detection in Multi-Source Multi-Spectral Earth Observation Images – AutoCloud”, Invitation to tender ESA/AD/1-8373/15/I-NB – “VAE: Next Generation EO-based Information Services”. doi:10.13140/RG.2.2.34162.71363. arXiv:1701.04256. Retrieved Jan., 8, 2017 from https://arxiv.org/ftp/arxiv/papers/1701/1701.04256.pdf

Baraldi, A., & Humber, M. (2013). Quality assessment of pre-classification maps generated from spaceborne/airborne multi-spectral images by the Satellite Image Automatic Mapper™ and Atmospheric/Topographic Correction™-Spectral Classification software products: Part 1 – Theory. IEEE Journal of Selected Topics Applications Earth Observation Remote Sensing, 3(3), 1307–1329.

Baraldi, A., Humber, M., & Boschetti, L. (2013). Quality assessment of pre-classification maps 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 Sensing, 5, 5209–5264. doi:10.3390/rs51035209

Baraldi, A., Puzzolo, V., Blonda, P., Bruzzone, L., & Tarantino, C. (2008). Automatic spectral rule-based preliminary mapping of calibrated Landsat TM and ETM+ images. IEEE Transactions on Geoscience and Remote Sensing, 44, 2563–2586. doi:10.1109/TGRS.2006.874140

Baraldi, A., & Soares, J. V. B. (2017). Multi-objective software suite of two-dimensional shape descriptors for object-based image analysis, subjects: computer vision and pattern recognition (cs.CV). arXiv:1701.01941. Retrieved January, 8, 2017 from https://arxiv.org/ftp/arxiv/papers/1701/1701.01941.pdf

Baraldi, A., Tiede, D., Sudmanns, M., Beligu, M., & Lang, S. (2016, September 14–16). Automated near real-time Earth observation Level 2 product generation for semantic querying. GeoBIA. Enschede, The Netherlands; University of Twente Faculty of Geo-Information and Earth Observation (ITC).

Baraldi, A., Wassenaar, T., & Kay, S. (2010). Operational performance of an automatic preliminary spectral rule-based decision-tree classifier of spaceborne very high resolution optical images. IEEE Transactions on Geoscience and Remote Sensing, 48, 3482–3502. doi:10.1109/TGRS.2010.2046741

Beauchemin, M., & Thomson, K. (1997). The evaluation of segmentation results and the overlapping area matrix. International Journal of Remote Sensing, 18, 3895–3899. doi:10.1080/0143116972176270

Belward, A. (Ed.). (1996). The IGBP-DIS global 1km land cover data set “DIScover” Proposal and implementation plans. IGBP-DIS working paper 13. Ispra, Varese, Italy: International Geosphere Biosphere Programme. European Commission Joint Research Center.

Benvenente, R., Vannrell, M., & Baldrich, R. (2008). Parametric fuzzy sets for automatic color naming. Journal of the Optical Society of America A, 25, 2582–2593. doi:10.1364/JOSAA.25.002582

Berlin, B., & Kay, P. (1969). Basic color terms: Their universality and evolution. Berkeley: University of California.

Benedus, P., & Noran, O. (2017). Data rich – but information poor. In L. Camarinho-Matos, H. Afsarmanesh, & R. Fornasiero (Eds), Collaboration in a Data-Rich World: PRO-VE 2017 (Vol. 506, pp. 206–214). IFIP Advances in Information and Communication Technology.

Bishop, C. M. (1995). Neural networks for pattern recognition. Oxford, UK: Clarendon.

Bishop, M. P., & Colby, J. D. (2002). Anisotropic reflectance correction of SPOT-3 HRV imagery. International Journal of Remote Sensing, 23(10), 2125–2131. doi:10.1080/01431160110097231

Bishop, M. P., Shroder, J. F., & Colby, J. D. (2003). Remote sensing and geomorphometry for studying relief production in high mountains. Geomorphology, 55(1–4), 345–361. doi:10.1016/S0169-555X(03)00149-1

Bishr, Y. (1998). Overcoming the semantic and other barriers to GIS interoperability. International Journal of Geographical Information Science, 12, 299–314. doi:10.1080/0144319982241806

Blaschke, T., Hay, G. J., Kelly, M., Hofmann, P., Addink, E., … Tiede, D. (2014). Geographic object-based image analysis - towards a new paradigm. ISPRS Journal of Photogrammetry and Remote Sensing, 87, 180–191. doi:10.1016/j.isprsjprs.2013.09.014

Boschetti, L., Fassò, S. P., & Brivio, P. A. (2004). Analysis of the conflict between omission and commission in low spatial resolution dichotomic thematic products: The Pareto boundary. Remote Sensing of Environment, 91, 280–292. doi:10.1016/j.rse.2004.02.015

Boschetti, L., Roy, D. P., Justice, C. O., & Humber, M. L. (2015). MODIS-Landsat fusion for large area 30 m burned area mapping. Remote Sensing of Environment, 161, 27–42. doi:10.1016/j.rse.2015.01.022

Bosward, M., Ferenec, J., & Otahel, J. (2000). CORINE land cover, technical guide – Addendum 2000 (Technical report No 40). European Environment Agency.

Capurro, R., & Hjelkran, B. (2003). The concept of information. Annual Review of Information Science and
Baraldi et al., Cogent Geoscience (2018), 4: 1467357
https://doi.org/10.1080/23312041.2018.1467357

Technology, 37, 343–411. doi:10.1002/arls.1440370109
Chavez, P. (1988). An improved dark-object subtraction technique for atmospheric smoothing correction of multispectral data. Remote Sensing of Environment, 24, 655–479. doi:10.1016/0034-4257(88)90199-3
Cherkassky, V., & Mulier, F. (1998). Learning from data: Concepts, theory, and methods. New York, NY: Wiley.
Cimpoi, M., Maji, S., Kokkinos, I., & Vedaldi, A. (2014). Deep filter banks for texture recognition, description, and segmentation. CoRR. abs/1411.6836.
CNES. (2015). Venus satellite sensor level 2 product. Retrieved January, 5, 2016 from https://venus.cnrs.fr/en/VENUS/prod_l2.htm
Congalton, R. G., & Green, K. (1999). Assessing the accuracy of remotely sensed data. Boca Raton, FL: Lewis Publishers.
Couclelis, H. (2010). Ontologies of geographic information. International Journal of Geographical Information Science, 24(12), 1785–1809. doi:10.1080/13658816.2010.484392
D’Elia, S. (2012). Personal communication. European Space Agency.
Despini, F., Teggi, S., & Baraldi, A. (2014, September 22). Methods and metrics for the assessment of pan-sharpening algorithms. In L. Bruzzone, J. A. Benediktsson, & F. Bovo, (Eds.), SPIE Proceedings, Vol. 9244: Image and Signal Processing for Remote Sensing XX. Amsterdam, Netherlands.
Deutsches Zentrum für Luft- und Raumfahrt e.V. (DLR) and VEGA Technologies. (2013). Sentinel-2 MSI – level 2A products algorithm theoretical basis document (Document S2PAD-ATBD-0001). European Space Agency.
DiCarlo, J. (2017). The science of natural intelligence: Reverse engineering primate visual perception. Keynote, CVPR17 Conference. Retrieved January, 30, 2018 from https://www.youtube.com/watch?v=ilbbVkhMgo
Di Gregorio, A., & Hansen, L. (2000). Land Cover Classification System (LCCS): Classification concepts and user manual. Rome, Italy: FAO, FAO Corporate Document Repository. Retrieved February, 10, 2015 from, http://www.fao.org/docrep/003/X0596E/X0596e00.htm
DigitalGlobe (2016). Automated Land Cover Classification. Retrieved November, 13, 2016 from http://gbdxdocs.digitalglobe.com/docs/automated-land-cover-classification
Dillencourt, M. B., Samet, H., & Tamminen, M. (1992). A general approach to connected component labeling for arbitrary image representations. Journal of the ACM, 39, 253–280. doi:10.1145/128749.128750
Dorigo, W., Richter, R., Baret, F., Bamber, R., & Wagner, W. (2009). Enhanced automated canopy characterization from hyperspectral data by a novel two step radiative transfer model inversion approach. Remote Sensing, 1, 1139–1170. doi:10.3390/rs1041139
Duke Center for Instructional Technology. (2016). Measurement: Process and outcome indicators. Retrieved June, 20, 2016 from http://patientsafetyed.dhhs.duke.edu/module_a/measurement/measurement.html
Dumitru, C. O., Cui, S., Schwanz, G., & Datcu, M. (2015). Information content of very-high-resolution SAR images: Semantics, geographic context, and ontologies. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 8(4), 1635–1650. doi:10.1109/JSTARS.2014.2363595
Elkan, C. (2003). Using the triangle inequality to accelerate k-means. International Conference Machine Learning.
Environmental Protection Agency (EPA). (2007). “Definitions” in Multi-Resolution Land Characteristics Consortium (MRLC). Retrieved November, 13, 2013 from http://www.epa.gov/mrlc/definitions.html#2001
Environmental Protection Agency (EPA). (2013). Western ecology division. Retrieved November, 13, 2013 from http://www.epa.gov/wed/pages/ecoregions.htm
Etzioni, O. (2017). “What shortcomings do you see with deep learning?” Quora. Retrieved January, 8, 2018 from https://quora.com/What-shortcomings-do-you-see-with-deep-learning
European Space Agency (ESA). (2015). Sentinel-2 user handbook, standard document, issue 1 rev 2.
Feldman, J. (2013). The neural binding problem(s). Cognition Neuroscience, 7, 1–11. doi:10.1007/s11571-012-9219-8
Feng, C.-C., & Flewelling, D. M. (2006). Assessment of semantic similarity between land use/land cover classification systems. Computers, Environment and Urban Systems, 28, 229–246. doi:10.1016/j.coeur.2015.03.0002-06
Foga, S., Scaramuzza, P., Guo, S., Zhu, Z., Dilley, R., Jr, Beckmann, T.,… Laue, B. (2017). Cloud detection algorithm comparison and validation for operational Landsat data products. Remote Sensing of Environment, 194, 379–390. doi:10.1016/j.rse.2017.03.026
Fanseca, F., Egenhofer, M., Agouris, P., & Camara, G. (2002). Using ontologies for integrated geographic information systems. Transactions in GIS, 6, 231–257. doi:10.1111/1467-9671.00109
Fowler, M. (2003). UML Distilled (3rd ed.). Boston, MS: Addison-Wesley.
Frintrop, S. (2011). Computational visual attention. In A. A. Salah & T. Gevers (Eds.), Computer analysis of human behavior, advances in pattern recognition. Springer.
Fritzke, B. (1997a). Some competitive learning methods. Retrieved March, 17, 2015 from http://www.denonig.de/JavaPaper17.html
Fritzke, B. (1997b). The LBG-U method for vector quantization - An improvement over LBG inspired from neural networks. Neural Processing Letters, 5(1), 35–45. doi:10.1023/A:1009653226428
GeoTerraImage. (2015). Provincial and national land cover 30m. Retrieved September, 22, 2015 from http://www.geoterraimage.com/products/landcover.php
Gevers, T., Glijsen, A., van de Weijer, J., & Geusebroek, J. M. (2012). Color in computer vision. Hoboken, Nj: Wiley.
Ghahramani, Z. (2011). Bayesian nonparametrics and the probabilistic approach to modelling. Philosophical Transactions R Social A, 1–27.
Ghamisi, P., Yokoya, N., Li, J., Liao, W., Liu, S., Plaza, J.,… Plaza, A. (2017). Advances in hyperspectral image and signal processing. IEEE Geoscience Remote Sensing Magazine, 12(37), 78.
Griffin, L. D. (2006). Optimality of the basic colour categories for classification. Journal of the Royal Society Interface, 3, 71–85. doi:10.1098/rsif.2005.0076
Griffith, G. E., & Omernik, J. M. (2009). Ecoregions of the United States-Level III (EPA). In C. J. Cleveland (Ed.), Encyclopedia of earth. Washington, DC: Environmental Information Coalition, National Council for Science and the Environment.
Group on Earth Observation (GEO). (2005). The Global Earth Observation System of Systems (GEOSS) 10-Year Implementation Plan, adopted 16 February 2005. Retrieved January, 10, 2012 from http://www.earthobservations.org/docs/10-Year%20Implementation%20Plan.pdf
architecture. Retrieved March, 8, 2015 from http://www.opengeospatial.org/standards/is
Ortiz, A., & Oliver, G. (2006). On the use of the overlapping area matrix for image segmentation evaluation: A survey and new performance measures. Pattern Recognition Letters, 27, 1916–1926. doi:10.1016/j.patrec.2006.05.002
Parisi, D. (1991). La scienza cognitiva tra intelligenza artificiale e vita artificiale. In Neuroscience e Scienze dell’Artificiale: Dal Neurone all’Intelligenza. Bologna, Italy: Petron Editore.
Patané, G., & Russo, M. (2001). The enhanced-LBG algorithm. Neural Networks, 14(9), 1219–1237. doi:10.1016/S0893-6080(01)00104-6
Patané, G., & Russo, M. (2002). Fully automatic clustering system. IEEE Transactions on Neural Networks, 13(6), 1185–1198. doi:10.1109/28.840226
Pearl, J. (2000). Causality: Models, reasoning and inference. New York, NY: Cambridge University Press.
Piaget, J. (1970). Genetic epistemology. New York, NY: Columbia University Press.
Pontius, R. G., Jr., & Connors, J. (2006, July 5–7). Expanding the conceptual, mathematical and practical methods for map comparison. In M. Gaetano & M. Painho (Eds). Proceedings of the 7th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences (pp. 64–79). Lisbon. Instituto Geográfico Português.
Pontius, R. G., Jr., &Millones, M. (2011). Death to Kappa: Birth of quantity disagreement and allocation disagreement for accuracy assessment. International Journal of Remote Sensing, 32(15), 4407–4429. doi:10.1080/01431161.2011.552923
Riano, D., Chuvieco, E., Salas, J., & Aguado, I. (2003). Assessment of different topographic corrections in landsat-TM data for mapping vegetation types (2003). IEEE Transactions on Geoscience and Remote Sensing, 41(5), 1056–1061. doi:10.1109/TGRS.2003.811693
Richter, R., & Schläpfer, D. (2012a). Atmospheric/topographic correction for satellite imagery – ATCOR-2/3 user guide, version 8.2 BETA. Retrieved April, 12, 2013 from http://www.dlr.de/oecl/Portal/60/Resources/dokumente/5_tech_mod/atcor3_manual_2012.pdf
Richter, R., & Schläpfer, D. (2012b). Atmospheric/Topographic correction for airborne imagery – ATCOR-4 User Guide, Version 6.2 BETA. Retrieved April, 2013 from http://www.dlr.de/oecl/Portal/60/Resources/dokumente/5_tech_mod/atcor4_manual_2012.pdf
Roy, D. P., Ju, J., Kline, K., Scaramuzza, P. L., Kovalsky, V., Hansen, M., … Zhang, C. (2010). Web-enabled Landsat Data (WELD): Landsat ETM+ composited mosaics of the conterminous United States. Remote Sensing of Environment, 114, 35–49. doi:10.1016/j.rse.2009.08.011
Salmon, B., Wessels, K., van den Bergh, F., Steenkamp, K., Kleynhans, W., Swanepeol, D., … Kovalsky, V. (2013, July 21–26). Evaluation of rule-based classifier for Landsat-based automated land cover mapping in Southern Africa. IEEE International Geoscience Remote Sensing Symposium (IGARSS) (pp. 4301–4304).
Schaepman-Strub, G., Schaepman, M. E., Painter, T. H., Dangel, S., & Martonchik, J. V. (2006). Reflectance quantities in optical remote sensing - Definitions and case studies. Remote Sensing of Environment, 103, 27–42. doi:10.1016/j.rse.2006.03.002
Schläpfer, D., Richter, R., & Hueni, A. (2009). Recent developments in operational atmospheric and radiometric correction of hyperspectral imagery, in Proc. 6th EARSeL SIG IS Workshop, 16–19 March 2009. Retrieved July, 14, 2012 from http://www.earselisht.tau.ac.il/e-earsel6/CDF/PDF/earsel-PROCEEDINGS/035%20shl%20poster.pdf
Shackelford, K., & Davis, C. H. (2003a). A combined fuzzy pixel-based and object-based approach for classification of high-resolution multispectral data over urban areas. IEEE Transactions on Geoscience and Remote Sensing, 41(10), 2354–2364. doi:10.1109/TGRS.2003.815972
Shackelford, K., & Davis, C. H. (2003b). A hierarchical fuzzy classification approach for high-resolution multi-spectral data over urban areas. IEEE Transactions on Geoscience and Remote Sensing, 41(9), 1920–1932. doi:10.1109/TGRS.2003.814627
Sheskin, D. (2000). Handbook of parametric and nonparametric statistical procedures. Boca Raton, FL: Chapman & Hall/CRC.
Shotton, J., Winn, J., Rother, C., & Criminisi, A. (2009). Texton boost for image understanding: Multi-class object recognition and segmentation by jointly modeling texture, layout, and context. International Journal of Computer Vision, 81(1), 2–23. doi:10.1007/s11263-009-0109-1
Simonetti, D., Marelli, A., & Eva, H. (2015b). Impact toolbox. (JRC Technical EUR 27358 EN).
Simonetti, D., Simonetti, E., Sztanoi, Z., Lupi, A., & Eva, H. D. (2015a). First results from the phenology based synthesis classifier using Landsat-8 imagery. IEEE Geoscience and Remote Sensing Letters, 12(7), 1496–1500. doi:10.1109/LGRS.2015.2409982
Smuelders, A., Worringer, M., Santini, S., Gupta, A., & Jain, R. (2000). Content-based image retrieval at the end of the early years. IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(12), 1349–1380. doi:10.1109/34.895972
Soares, J., Baraldi, A., & Jacobs, D. (2014). Segment-based simple-connectivity measure design and implementation (Tech. Rep). College Park, MD: University of Maryland.
Sonka, M., Hlavac, V., & Boyle, R. (1994). Image processing, analysis and machine vision. London, UK: Chapman & Hall.
Sowa, J. F. (2000). Knowledge representation: Logical, Philosophical, and Computational Foundations. Pacific Grove, CA: Brooks/Cole.
Stehman, S. V. (1999). Comparing thematic maps based on map value. International Journal of Remote Sensing, 20, 2347–2366. doi:10.1080/014311619901431169212065
Stehman, S. V., & Czaplewski, R. L. (1998). Design and analysis for thematic map accuracy assessment: Fundamental principles. Remote Sensing of Environment, 64, 331–344. doi:10.1016/S0034-4257(98)00010-8
Swain, P. H., & Davis, S. M. (1978). Remote sensing: The quantitative approach. New York, NY: McGraw-Hill.
Tiede, D., Baraldi, A., Sudmanns, M., Belgu, M., & Lang, S. (2016, March 15–17). ImageQuerying (IQ) – Earth observation image content extraction & querying across time and space. Submitted (Oral presentation and poster session), ESA 2016 Conf. on Big Data From Space, BIDS ’16. Santa Cruz de Tenerife, Spain.
Trimble. (2015). eCognition® Developer 9.0 Reference Book.
Tsotsos, J. K. (1990). Analyzing vision at the complexity level. Behavioral and Brain Sciences, 13, 423–445. doi:10.1017/S0140525X107007977
Vermote, E., & Saleous, N. (2007). LEDAPS surface reflectance product description - Version 2.0. University of
Vogelmann, J. E., Howard, S. M., Yang, L., Larson, C. R., Wylie, B. K., & van Driel, N. (2001). Completion of the 1990s National Land Cover Data set for the conterminous United States from Landsat Thematic Mapper data and ancillary data sources. Photo Engineering Remote Sensing, 67, 650–662.

Vogelmann, J. E., Sohl, T. L., Campbell, P. V., & Shaw, D. M. (1998). Regional land cover characterization using Landsat Thematic Mapper data and ancillary data sources. Environmental Monitoring and Assessment, 51, 415–428. doi:10.1007/A:1005996900217

Web-Enabled Landsat Data (WELD) Tile FTP. 2015. Retrieved December, 12, 2016 from https://weld.cr.usgs.gov/

Wenwen, L. M., Goodchild, F., & Church, R. L. (2013). An efficient measure of compactness for 2D shapes and its application in regionalization problems. International Journal of Geographical Info Science, 1–24.

Wickham, J. D., Stehman, S. V., Fry, J. A., Smith, J. H., & Homer, C. G. (2010). Thematic accuracy of the NLCD 2001 land cover for the conterminous United States. Remote Sensing of Environment, 114, 1286–1296. doi:10.1016/j.rse.2010.01.018

Yan, C., Huang, Q., Li, Z., Liu, K., & Hu, F. (2017). Big Data and cloud computing: Innovation opportunities and challenges. International Journal of Digital Earth, 10(1), 13–53. doi:10.1080/17538947.2016.1239771

Zadeh, L. A. (1965). Fuzzy sets. Information and Control, 8, 338–353. doi:10.1016/S0019-9958(65)90241-X