International Journal of *Geo-Information*

MDPI

Article

AutoCloud+, a "Universal" Physical and Statistical Model-Based 2D Spatial Topology-Preserving Software for Cloud/Cloud–Shadow Detection in Multi-Sensor Single-Date Earth Observation Multi-Spectral Imagery—Part 2: Outcome and Process Requirements Specification, Information/Knowledge Representation, System Design, Algorithm, Implementation and Preliminary Experimental Results

Andrea Baraldi [1,2,*] and Dirk Tiede [1]

1 Department of Geoinformatics – Z_GIS, University of Salzburg, Schillerstr. 30, 5020 Salzburg, Austria; dirk.tiede@sbg.ac.at
2 Italian Space Agency (ASI), Via del Politecnico, 00133 Rome RM, Italy.
* Correspondence: andrea6311@gmail.com; Tel.: +39-329-2648110

Received: date; Accepted: date; Published: date

Abstract: The European Space Agency (ESA) defines Earth observation (EO) Level 2 information product the stack of: (i) a single-date multi-spectral (MS) image, radiometrically corrected for atmospheric, adjacency and topographic effects, with (ii) its data-derived scene classification map (SCM), whose thematic map legend includes quality layers cloud and cloud–shadow. Never accomplished to date in an operating mode by any EO data provider at the ground segment, systematic ESA EO Level 2 product generation is an inherently ill-posed computer vision (CV) problem (*chicken-and-egg* dilemma) in the multi-disciplinary domain of cognitive science, encompassing CV as *subset-of* artificial intelligence (AI). This research and technological development (RTD) study aims at creating a "universal" AutoCloud+ software system in operating mode, capable of systematic cloud and cloud–shadow quality layers detection in multi-sensor, multi-temporal and multi-angular EO *big data cubes* characterized by the five Vs, namely, volume, variety, veracity, velocity and value. For the sake of readability this paper is divided in two. The previous Part 1 highlights why AutoCloud+ is important in a broad context of systematic ESA EO Level 2 product generation at the ground segment within a "seamless chain of innovation" needed for a new era of Space Economy 4.0. In the notion of Space 4.0, the ESA EO Level 2 product definition is proposed as the new standard of EO Analysis Ready Data (ARD) format. In the present Part 2 (considered as Supplementary Materials), first, the proposed "universal" AutoCloud+ software system is instantiated in terms of outcome and process requirements specification, information/knowledge representation, system design, algorithm and implementation. Second, preliminary experimental results collected from the AutoCloud+ software prototype are presented and discussed in comparison with those of standard cloud/cloud–shadow detectors, available either open source or free of cost, such as the free-of-cost single-date sensor-specific ESA Sen2Cor software toolbox, to be run on the user side, and the multi-date multi-sensor MAJA software, developed and run by CNES/ CESBIO/ DLR.

Keywords: artificial intelligence; color naming; color constancy; cognitive science; computer vision; object-based image analysis (OBIA); physical and statistical data models; radiometric calibration; semantic content-based image retrieval; spatial topological and spatial non-topological information components.

1. Introduction

To contribute toward filling an analytic and pragmatic information gap from Earth observation (EO) *big data*, characterized by the five Vs of volume, variety, veracity, velocity and value [1], to timely, comprehensive and operational EO data-derived value-adding products and services (VAPS), this paper presents a research and technological development (RTD) study of a "universal" AutoCloud+ computer vision (CV) software system for cloud and cloud–shadow quality layer detection in multi-sensor, multi-temporal and multi-angular EO *big data cubes*, in compliance with the intergovernmental Group on Earth Observations (GEO)-Committee on Earth Observation Satellites (CEOS) Quality Accuracy Framework for Earth Observation (QA4EO) Calibration/Validation (*Cal/Val*) requirements [2] and with the visionary goal of a GEO's implementation plan for years 2005-2015 of a Global Earth Observation System of Systems (GEOSS) [3], unaccomplished to date.

For the sake of readability this paper is divided in two. The previous Part 1 highlights why AutoCloud+ is important in a broad context of systematic European Space Agency (ESA) EO Level 2 information product generation at the ground segment, aimed at harmonization of missions acquiring multi-source EO data across time and geographic space, within a "seamless chain of innovation" needed for a new era of Space Economy 4.0 [4] (refer to Section 2 in the Part 1). In the notion of Space 4.0 (see Figure 8 in the Part 1), ESA EO Level 2 product was regarded as an "augmented" standard of EO Analysis Ready Data (ARD), alternative to existing U.S. Landsat ARD [5–9] and CEOS ARD for Land (CARD4L) format definitions [10]. In the present Part 2 (proposed as Supplementary Materials), first, a "universal" AutoCloud+ CV software system is instantiated at the Marr five levels of understanding of an information processing system, specifically [11–15]:

- outcome and process requirements specification, including computational complexity estimation,
- information/knowledge representation,
- system design (architecture),
- algorithm, and
- implementation.

Among these five levels, the three more abstract ones, namely, outcome and process requirements specification, information/knowledge representation and system design, are typically considered the linchpin of success of an information processing system, rather than algorithm and implementation [11–15]. Second, the AutoCloud+ prototypical implementation is compared with standard cloud/cloud–shadow detectors available either open source or free of cost (see Table 5 in the Part 1), such as the single-date sensor-specific ESA Sentinel 2 (atmospheric and topographic) Correction Prototype Processor (Sen2Cor), to be run free-of-cost on the user side [16,17] (see Figure 11 in the Part 1), and the multi-date Multisensor Atmospheric Correction and Cloud Screening (MACCS)-Atmospheric/Topographic Correction (ATCOR) Joint Algorithm (MAJA), developed and run by the Centre national d'études spatiales (CNES)/ Centre d'Etudes Spatiales de la Biosphère (CESBIO)/ Deutsches Zentrum für Luft- und Raumfahrt (German Aerospace Center, DLR) [18–20], which incorporates capabilities of the ATCOR commercial software toolbox [21–24], refer to Section 3 in the Part 1.

The rest of the present Part 2 is organized as follows. To make this Part 2 self-contained, background definitions and concepts are reported in Section 2. In Sections 3 and 4 the

AutoCloud+ CV system software instantiation is proposed in terms of methods and materials respectively. Preliminary experimental results are presented in Section 5 and discussed in Section 6. Conclusions are reported in Section 7.

2. Background definitions and concepts

In recent years the European Space Agency (ESA) has been defining an ESA Earth observation (EO) Level 2 information product as follows [16,17] (refer to Section 1 in Part 1):

(i) a single-date multi-spectral (MS) image, radiometrically corrected for atmospheric, adjacency and topographic effects,
(ii) stacked with its data-derived scene classification map (SCM), whose general-purpose, user- and application-independent thematic map legend includes quality layers cloud and cloud–shadow,
(iii) to be systematically generated at the ground segment, automatically (without human–machine interaction) and in near real-time.

Never accomplished to date in an operating mode by any EO data provider at the ground segment, systematic ESA EO Level 2 product generation is an inherently ill-posed computer vision (CV) problem (*chicken-and-egg* dilemma) [11,13,25,26] in the multi-disciplinary domain of cognitive science [27–32] (see Figure 2 in the Part 1), encompassing CV as *subset-of* artificial (general) intelligence (AI) [27–32], i.e., '[AI ⊃ CV] → cognitive science' in symbols of the standard Unified Modeling Language (UML) for graphical modeling of object-oriented software [33], where symbol '→' denotes relationship *part-of* pointing from the supplier to the client, not to be confused with relationship *subset-of*, '⊃', meaning specialization with inheritance from the superset to the subset.

Encompassing both biological vision and CV in the cognitive science domain (see Figure 2 in the Part 1), the word *vision* is synonym for inherently ill-posed *scene-from-image reconstruction and understanding* [11,13,25,26], see Figure 6 in the Part 1. Vision is a cognitive *(information-as-data-interpretation*) problem [27] very difficult to solve because (refer to Section 2 in the Part 1): (i) non-polynomial (NP)-hard in computational complexity [34,35], (ii) inherently ill-posed in the Hadamard sense [36], i.e., vision admits no solution, multiple solutions or, if the solution exists, the solution's behavior changes continuously with the initial conditions [25,26]. Vision is inherently ill-posed because affected by: (I) a 4D-to-2D data dimensionality reduction from the 4D geospatial-temporal scene-domain to the (2D, planar) image-domain, e.g., responsible of occlusion phenomena, and (II) a semantic information gap from ever-varying sub-symbolic sensory data (sensations) in the physical world to stable symbolic percepts in the mental model of the physical world (modeled world, world ontology, real-world model) [11,12,25,27,37–40]. Since it is inherently ill-posed, vision requires *a priori* knowledge in addition to sensory data to become better posed for numerical solution [41,42].

Largely oversighted by the remote sensing (RS) and CV literature, an undisputable true fact (observation) is that, in general, spatial information dominates color information in vision [11,25]. This commonsense knowledge is obvious, but not trivial. On the one hand, it may sound awkward to many readers, including RS experts and CV practitioners. On the other hand, it is acknowledged implicitly by all human beings wearing sunglasses: human panchromatic vision is nearly as effective as chromatic vision in scene-from-image reconstruction and understanding [11]. This true fact means that spatial information dominates both the 4D geospatial-temporal scene-domain and the (2D, planar) image-domain involved with the cognitive task of vision, see Figure 6 in the Part 1. This evidence is acknowledged by the Tobler's first law (TFL) of geography, familiar to geographers working in the real-world (geographic) domain. The TFL of geography states that "all things are related, but nearby things are more related than distant things"[43], although certain phenomena clearly constitute exceptions [44]. Obscure to many geographers familiar with the TFL formulation, the statistical concept of spatial autocorrelation is the quantitative counterpart of the qualitative TFL of geography [11]. The relevance of spatial autocorrelation in both the 4D geospatial-temporal scene-domain and the (2D) image-domain involved with vision is at the very foundation of the object-based image analysis (OBIA) approach to CV, originally conceived around year 2000 by the geographic information science (GIScience) community as a viable alternative to traditional 1D spatial-context insensitive (pixel-based) image analysis [45,46].

In more detail, the observation that, in general, spatial information dominates color information in vision [11,25], proved by perceptual evidence about vision in primates [11,13,35,47–63], means that primary spatial topological information (e.g., adjacency, inclusion, etc.) [11,35,47,55,64] and spatial non-topological information (e.g., spatial distance, angle measure) components dominate secondary color information [25], which is the sole information available at the imaging sensor's spatial resolution, i.e., at the pixel level of spatial analysis (refer to Section 2 in the Part 1). Irrespective of this undisputable true fact (observation), to date, a large majority of EO image understanding (EO-IU) $\subset$ CV systems consists of 1D image analysis algorithms (see Figure 15 in the Part 1), either pixel-based, synonym for spatial context-insensitive and spatial topology non-preserving (non-retinotopic), or spatial context-sensitive (e.g., image object-based or local window-based), but spatial topology non-preserving (non-retinotopic) [11,35,47,55,64]. Intuitively, 1D image analysis algorithms are invariant to permutations in the 1D vector data sequence generated from a (2D) image [65], where image is synonym for 2D gridded data set (see Figure 15 in the Part 1). In short, 1D analysis of (2D) imagery is affected by a loss in data dimensionality. Inferred from Table 5 in the Part 1, instances of 1D image analysis algorithms are the popular CV systems for cloud and cloud–shadow quality layers detection available either open source or free-of-cost, such as the single-date multi-sensor Function of Mask (FMask) open source algorithm [66,67], the single-date single-sensor ESA Sen2Cor prototype processor, to be run free-of-cost on the user side [16,17] (see Figure 11 in the Part 1), and the multi-date multi-sensor MAJA, developed and run by CNES/ CESBIO/ DLR [18–20].

Alternative to 1D image analysis is 2D image analysis, which is spatial context-sensitive and spatial topology-preserving (retinotopic) [11,35,47,55,64], i.e., it is sensitive to permutations in the order of presentation of the input 2D data set [65] (see Figure 16 in the Part 1). In our understanding, 2D spatial topology-preserving mapping is the fundamental basis of success of multi-scale 2D spatial filter banks for image analysis (encoding, decomposition), synthesis (decoding, reconstruction) and classification (understanding), either deductive/physical model-based [11,13,35,47–63] or end-to-end inductive learning-from-data, such as increasingly popular deep convolutional neural networks (DCNNs) [65,68–71].

3. Methods

To overcome structural drawbacks of 1D image analysis (refer to Section 2 in the Part 1) and well-known failure modes of standard CV systems available open source or free-of-cost for cloud and cloud–shadow detection in MS imagery [19,72–75], such as the single-date open source FMask algorithm [66,67], the single-date sensor-specific ESA Sen2Cor software, to be run free-of-cost on the user side [16,17,72], and the multi-date multi-sensor MAJA algorithm, developed and run by CNES/ CESBIO/ DLR [18–20], which are all 1D image analysis approaches (refer to Table 5 in the Part 1), an inherently ill-posed "universal" AutoCloud+ CV system [11,76,77] was constrained as follows, to become better-posed for numerical solution [41,42], at the Marr five levels of understanding of an information processing system (refer to Section 1) [11–15].

- Outcome and process requirements specification: Suitable for cloud and cloud–shadow quality layer detection in MS imagery under operational constraints of "universality". To be considered "universal" in operating mode (refer to Section 2 in the Part 1), AutoCloud+ was required to be: (i) "fully automated", i.e., to run, it requires no human–machine interaction and no labeled data set for supervised inductive learning-from-data, (ii) near real-time, e.g., its computational complexity increases linearly with image size, (iii) robust to changes in input data acquired across space-time and sensors, (iv) robust to changes in MS imaging sensor's spatial and spectral resolution specifications, and (v) robust to changes in radiometric *Cal* metadata parameters, i.e., AutoCloud+ can be input with multi-source, multi-angular and multi-temporal MS imagery whether or not radiometrically calibrated into TOARF, SURF or surface albedo values. In other words, a "universal" AutoCloud+ software can be input with any MS imagery acquired by multiple platforms, either spaceborne or airborne, including unmanned aerial vehicles (UAVs), even when a lightweight optical imaging sensor, such as those mounted aboard small satellites [78] or small UAVs [79], does not feature any on-board radiometric calibrator to provide an acquired image with its image-specific radiometric *Cal* metadata file(s).

- Information/knowledge representation and system design: Hybrid (combined deductive and inductive) inference, where deductive and inductive inference alternate at increasing levels of specialization, see Figure 12 in the Part 1, to take advantage of each and overcome their shortcomings [11,14,15,80], refer to Section 2 in the Part 1.
- Information/knowledge representation and system design: 2D (retinotopic, spatial topology-preserving) image analysis, see Figure 16 in the Part 1, where primary spatial information and secondary color information in vision are combined according to a convergence-of-evidence approach [25], consistent with human symbolic reasoning, where eventually weak, but independent sources of evidence can be combined to infer strong conjectures, and where it is acknowledged that "vision goes symbolic almost immediately" [11,13].
- Information/knowledge representation and system design: Provided with feedback loops (feedback system) [35,47–52,55–57]. A hybrid feedback CV system is alternative to inductive feedforward CV systems, such as increasingly popular DCNNs trained from data end-to-end [65,68–71], deductive feedforward CV systems, such as the ESA Sen2Cor prototype processor, see Figure 11 and Table 5 in the Part 1, and hybrid feedforward CV systems, such as the "augmented" ATCOR software toolbox [24], see Figure 10 in the Part 1.
- Information/knowledge representation: the AutoCloud+ CV subsystem adopts a convergence-of-evidence approach [25], equivalent to a Bayesian naïve classifier [11,41,42]. According to the Bayesian law, a naive Bayes classifier assumes the "naive" conditional independence of input features F_i, i = 1, …, I, hence,

$$p(c|F_i, \dots, F_I) = p(c)\prod_{i=1}^{I} p(F_i|c). \quad (1)$$

Deeply investigated by the CV community [12,13,43] and by those portions of the RS community involved with traditional content-based image retrieval (CBIR) [81,82] and with EO-IU applications for biophysical variable estimation at the Earth surface [11,14,15,25,83–88], well-known visual features are: (i) color values, typically discretized by humans into a finite and discrete vocabulary of basic color (BC) names [89,90]; (ii) planar shape [11,86–92]; (iii) texture, defined as the perceptual spatial grouping of texture elements known as texels [11,12,93–95] or tokens [13]; (iv) inter-object spatial topological relationships, e.g., adjacency, inclusion, etc., and (v) inter-object spatial non-topological relationships, e.g., spatial distance, angle measure, etc. [11,12,25,37–39]. In vision, color is the sole visual property available at the imaging sensor's spatial resolution, i.e., at the pixel level of spatial analysis. In other words, pixel-based information is spatial context-independent, i.e., per-pixel information is exclusively related to color properties. Among the aforementioned visual variables, per-pixel color values are the sole non-spatial (spatial context-insensitive) numeric variable. It is easy to prove that, irrespective of their Pearson inter-feature cross-correlation, if any, individual sources of visual evidence, specifically, color, local shape, texture and inter-object spatial relationships, are statistically independent because, in general, Pearson's linear cross-correlation does not imply causation [11,12,42,96–98]. In the CV domain, if priors are ignored because considered equiprobable in a maximum class-conditional likelihood inference approach, alternative to a maximum *a posteriori* optimization criterion, and if a canonical interpretation based on frequentist statistics can be relaxed by fuzzy logic [99] into a membership function m(·) belonging to range 0.0–1.0, then the naive Bayes classifier, see Equation (1), becomes [11,83,84]:

$$p(c \mid \text{ColorValue}(x), \text{ShapeValue}(x), \text{TextureValue}(x), \text{SpatialRelationships}(x, \text{Neigh}(x))) =$$
$$p(c|F_i, \ldots, F_{I=4}) = p(c) \prod_{i=1}^{I=4} p(F_i|c) \propto$$
$$m(c \mid \text{ColorValue}(x), \text{ShapeValue}(x), \text{TextureValue}(x), \text{SpatialRelationships}(x, \text{Neigh}(x))) \propto$$
$$\min\{ \sum_{\text{ColorName}=1}^{\text{ColorVocabularyCardinality}} m_1(\text{ColorValue(x)}|\text{ColorName}) \cdot m_2(\text{ColorName}|c),\ m_3(\text{ShapeValue}(x) \mid c),\ m_4(\text{TextureValue}(x) \mid c),\ m_5(\text{SpatialRelationships}(x, \text{Neigh}(x)) \mid c)\} = \quad (2)$$
$$\min\{ m_2(\text{ColorName}^* \mid c),\ m_3(\text{ShapeValue}(x) \mid c),\ m_4(\text{TextureValue}(x) \mid c),\ m_5(\text{SpatialRelationships}(x, \text{Neigh}(x)) \mid c)\},\ c = 1, \ldots, \text{ObjectClassLegendCardinality},$$

where ColorName* ∈ {1, ColorVocabularyCardinality}, such that m_1(ColorValue(x)∣ ColorName*) = 1 and m_2(ColorName*∣ c) ∈ {0, 1},

where x is a spatial unit in the (2D) image-domain, either 0D point, 1D line or 2D polygon [100], Neigh(x) is a generic 2D spatial neighborhood of spatial unit x, colorValue(x) belongs to a MS measurement space $\Re^{MS}$, i.e., ColorValue(x) ∈ $\Re^{MS}$, and color space $\Re^{MS}$ is partitioned into a set of mutually exclusive and totally exhaustive hyperpolyhedra, equivalent to a discrete and finite vocabulary (categorical variable) of static color names, with ColorName = 1, .., ColorVocabularityCardinality. About Equation (2), the following considerations hold.

- Each numeric ColorValue(x) is a vector data equivalent to one point in a MS color space $\Re^{MS}$. In the MS color space, each point belongs to a single color name (hyperpolyhedron), identified as ColorName* in the static color vocabulary, i.e., ∀ ColorValue(x) ∈ $\Re^{MS}$, then equality $\sum_{\text{ColorName}=1}^{\text{ColorVocabularyCardinality}} m_1(\text{ColorValue(x)}|\text{ColorName}) = m_1(\text{ColorValue(x)}|\text{ColorName}^*) = 1$ holds, where m_1(ColorValue(x)∣ ColorName) ∈ {0, 1} is a binary (crisp) membership function, with ColorName = 1, …, ColorVocabularyCardinality.
- Equation (2) 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 [34].
- In Equation (2), a vocabulary of color names, physically equivalent to a partition of a numeric color hyperspace $\Re^{MS}$ into color name-specific hyperpolyhedra, is conceptually equivalent to a latent/ hidden/ hypothetical categorical variable, see Figure 1. In statistics, the popular concept of latent/hidden categorical variable was introduced to fill the information gap from input numeric observables, e.g., sub-symbolic sensory data such as color values in a color space, to an output categorical variable, such as a discrete and finite dictionary of LC classes in the mental model of the physical world (world ontology, world model). Latent/hidden variables are categorical variables not directly measured, but inferred from lower level variables. "The terms hypothetical variable or hypothetical construct may be used when latent variables correspond to abstract concepts, like perceptual categories or discrete mental states" [11,12,101,102]. Hence, to fill the semantic gap from input sub-symbolic sensory data to an output categorical variable of symbolic quality, an hypothetical categorical variable, such as a discrete and finite dictionary of BC names [89,90], is considered to be of "semi-symbolic" quality, i.e., superior to zero semantics, but inferior to the symbolic quality of the output categorical variable.

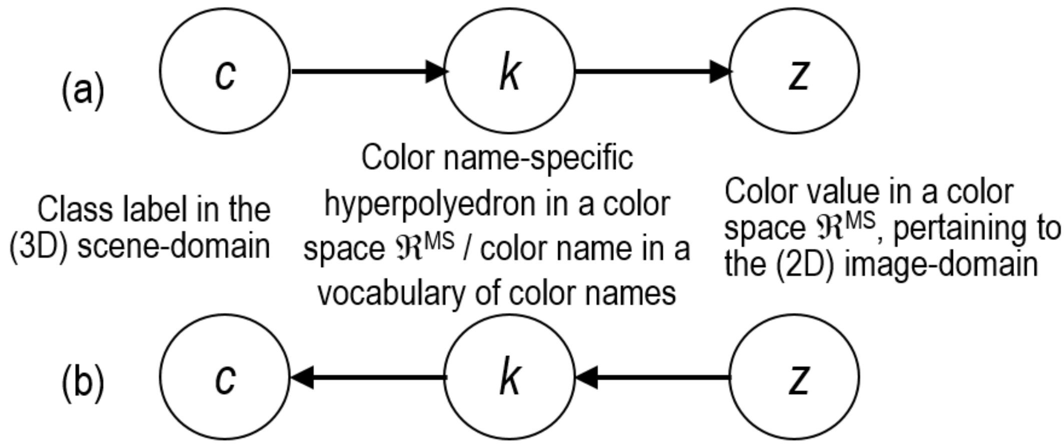


Figure 1. Graphical model of color naming, adapted from [101]. Let us consider z as a (sub-symbolic) numeric variable, such as MS color values of a population of spatial units, with vector data $z \in \Re^{MS}$, where $\Re^{MS}$ represents the MS data space, while c represents a categorical variable of symbolic classes in the physical world, with c = 1, …, ObjectClassLegendCardinality. (a) According to Bayesian theory, posterior probability p(c|z) $\propto$ p(z|c)p(c) = p(c) $\sum_{\mathrm{ColorName}=k=1}^{\mathrm{ColorVocabularyCardinality}} p(z|k)p(k|c)$, where color names, equivalent to color hyperpolyhedra in a numeric color space $\Re^{MS}$, provide a partition of the domain of change, $\Re^{MS}$, of numeric variable z. (b) For discriminative inference, the arrows in the graphical model are reversed using Bayes rule. Hence, a vocabulary of color names, physically equivalent to a partition of a numeric color space $\Re^{MS}$ into color name-specific hyperpolyhedra, is conceptually equivalent to a latent/ hidden/ hypothetical variable linking observables (sub-symbolic sensory data) in the real world, specifically, color values, to a categorical variable of semantic (symbolic) quality in the mental model of the physical world (world ontology, world model).

- Equation (2) shows that for any spatial unit x in the image-domain, either 0D point, 1D line or 2D polygon [100], 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_1(ColorValue(x)| ColorName*) = 1 holds, if condition m_2(ColorName*| c) = 0 is true according to a community-agreed binary relationship R: VocabularyOfColorNames $\Rightarrow$ LegendOfObjectClassNames (and vice versa) known *a priori* at first stage, see Table 6 in the Part 1, then the output membership function 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 (2) shows that first-stage static color naming of any spatial unit x in the image-domain, either (0D) pixel, (1D) line or (2D) polygon [100], allows the color-based stratification of unconditional multivariate spatial variables into color class-conditional multivariate data distributions, in agreement with the statistic stratification principle [103] and the divide-and-conquer (*dividi-et-impera*) problem solving approach [41,42,104]. 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" [103].
- System design (architecture), algorithm and implementation: The AutoCloud+ CV system requirements specification listed above is ambitious, but realistic, because AutoCloud+ can

rely upon several low-level CV ⊃ EO-IU software functions (modules) already implemented, tested (by their authors) and validated (by independent third-parties) in recent years [11,83,84,105]. These existing CV ⊃ EO-IU software units, capable of low-level (pre-attentive) vision tasks [11–13,35,47-52,54–63], can be combined according to an original six-stage hybrid feedback EO-IU system architecture, hereafter identified as QuickMap™ technology [13], see Figure 2, in compliance with the engineering principles of modularity, hierarchy and regularity considered mandatory by structured system design to guarantee scalability [104]. The six-stage hybrid feedback CV ⊃ EO-IU system shown in Figure 2 constitutes the core of the inference engine required by the closed-loop EO-IU4SQ system to systematically provide sub-symbolic EO *big data cubes* with meanings (semantics, intelligence), eligible for transforming a traditional Data and Information Access Services (DIAS), typically affected by a so-called Data-Rich Information-Poor (DRIP) syndrome, into an innovative AI4DIAS framework, see Figure 5 in the Part 1 [13, 106– 110].

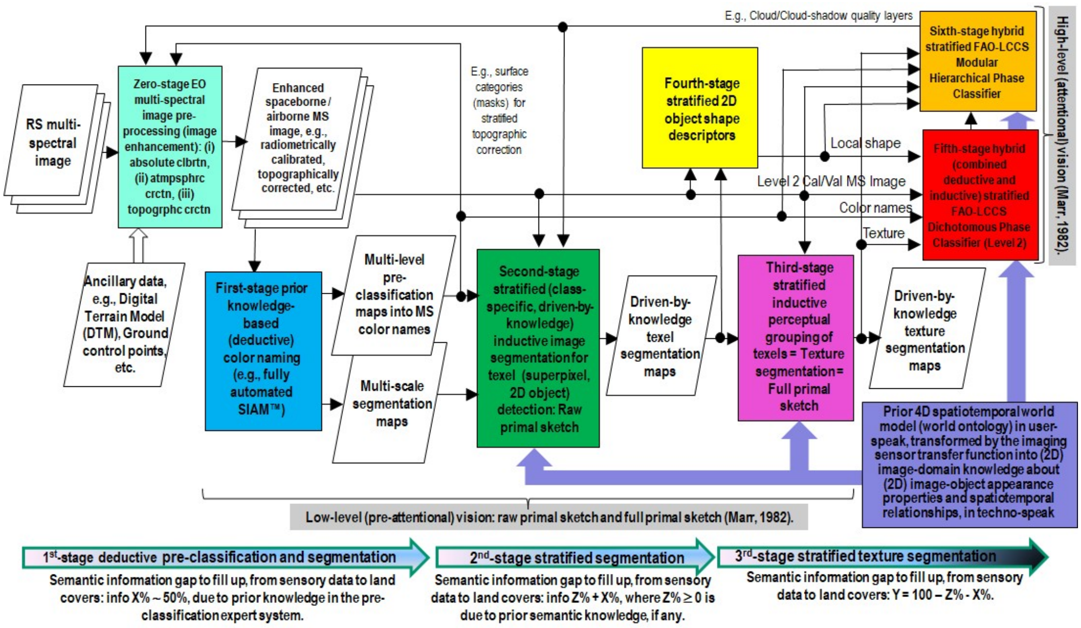


Figure 2. Six-stage hybrid (combined deductive and inductive) feedback EO image understanding (EO-IU) system design, identified as QuickMap™ technology, where acronym SIAM stays for Satellite Image Automatic Mapper (SIAM), a lightweight computer program for MS reflectance space hyperpolyhedralization into a static vocabulary of MS color names, superpixel detection and vector quantization (VQ) quality assessment [13–15,80,83,84,111–116]. The proposed six-stage hybrid EO-IU system architecture is based on a convergence-of-evidence approach to vision [25], consistent with Bayesian naïve classification [11,41,42], refer to Equation (2). Alternative to inductive feedforward EO-IU system architectures adopted by the RS mainstream, such as Deep Convolutional Neural Networks (DCNNs) trained from data end-to-end [65,68–71], the proposed six-stage hybrid EO-IU system design complies with the engineering principles of modularity, hierarchy and regularity considered necessary for scalability in structured system design [104]. Its hierarchy comprises a first-stage general-purpose, sensor-, application- and user-independent EO image understanding (classification) subsystem, followed by a second-stage sensor-, application- and user-specific EO image understanding subsystem. This two-stage EO-IU system design is fully consistent with the standard two-stage fully-nested Land Cover Classification System (LCCS) taxonomy promoted by the Food and Agriculture Organization (FAO) of the United Nations, where a first-stage 3-level 8-class Dichotomous Phase (DP) is preliminary to a second-stage Modular Hierarchical Phase (MHP) [117]. For the sake of visualization, each of the six EO data processing stages plus stage-zero for EO data pre-processing (enhancement) is depicted as a

rectangle with a different color fill. Visual evidence stems from multiple information sources, specifically, numeric color values quantized into categorical color names, local shape, texture and inter-object spatial relationships, either topological or non-topological. An example of first-stage general-purpose, user- and application-independent EO image classification taxonomy required by an ESA EO Level 2 Scene Classification Map (SCM) product is the 3-level 8-class FAO LCCS-DP legend, in addition to quality layers cloud and cloud–shadow. Second-stage EO image classification is user- and application-specific, where an SCM product of Level 3 or superior is provided with a map legend consistent with the FAO LCCS-MHP taxonomy [117], see Figure 1 in the Part 1.

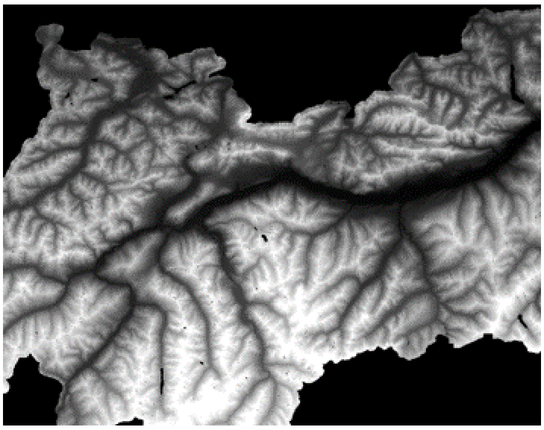
In more detail, at the two levels of understanding known as algorithm and implementation (refer to Section 1), the six-stage hybrid feedback CV ⊃ EO-IU system architecture shown in Figure 2 can benefit from an ensemble (library) of existing CV software functions (modules).

(i) Zero-stage EO image pre-processing (enhancement), to guarantee multi-source multi-temporal and multi-angular EO image harmonization and interoperability [11].

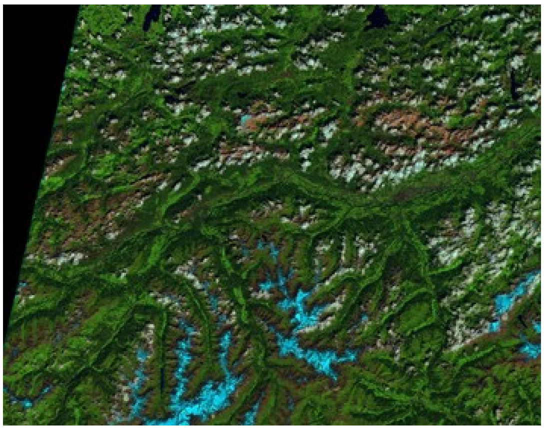
   I. Required input data set: MS image provided with a radiometric *Cal* metadata file, in agreement with the GEO-CEOS QA4EO *Cal* requirements [2].

      a. Absolute radiometric *Cal* of DNs into TOARF values, based on radiometric *Cal* gain and offset metadata parameters available per spectral channel.

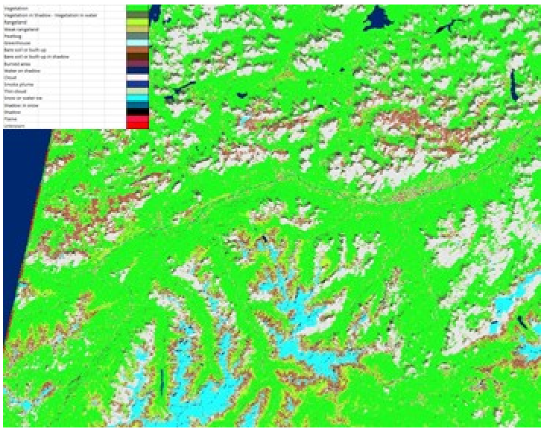
      b. Automated (without human–machine interaction) stratified (class-conditional, masked) atmospheric correction, stratified adjacency correction and stratified topographic correction (STRATCOR) of TOARF into SURF values at increasing levels of radiometric quality for data harmonization (reconciliation) purposes, see Figure 3, e.g., refer to [11,80].



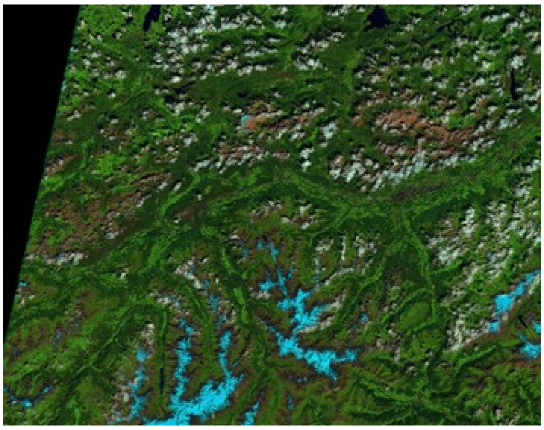
(a)



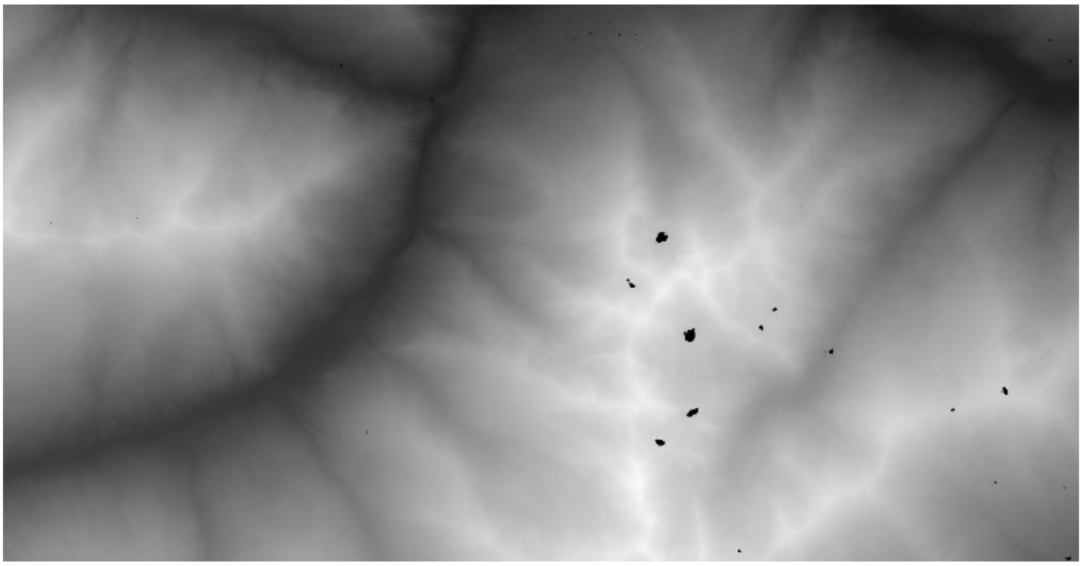
(b)



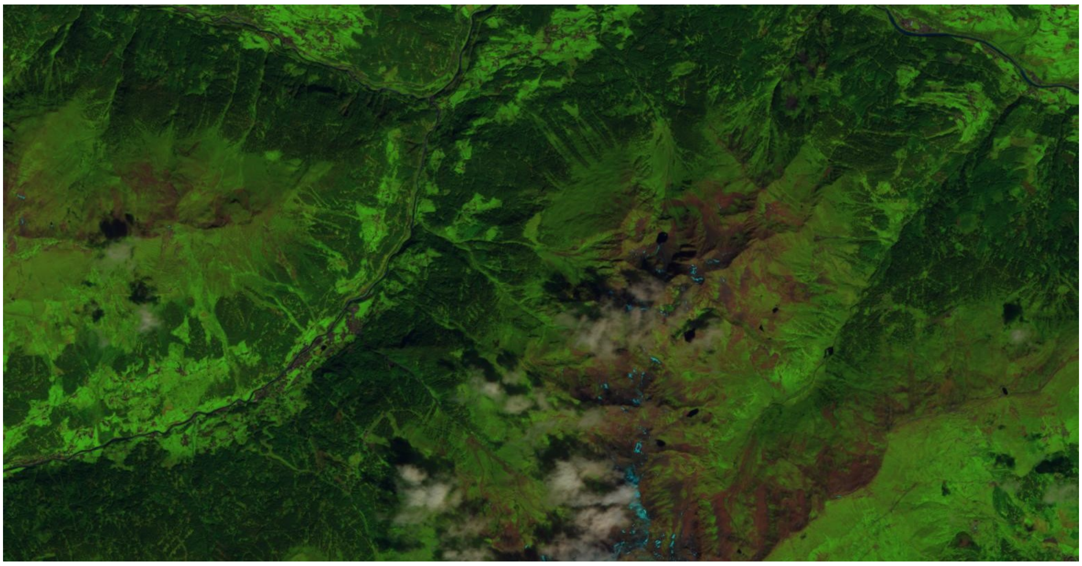
(c)



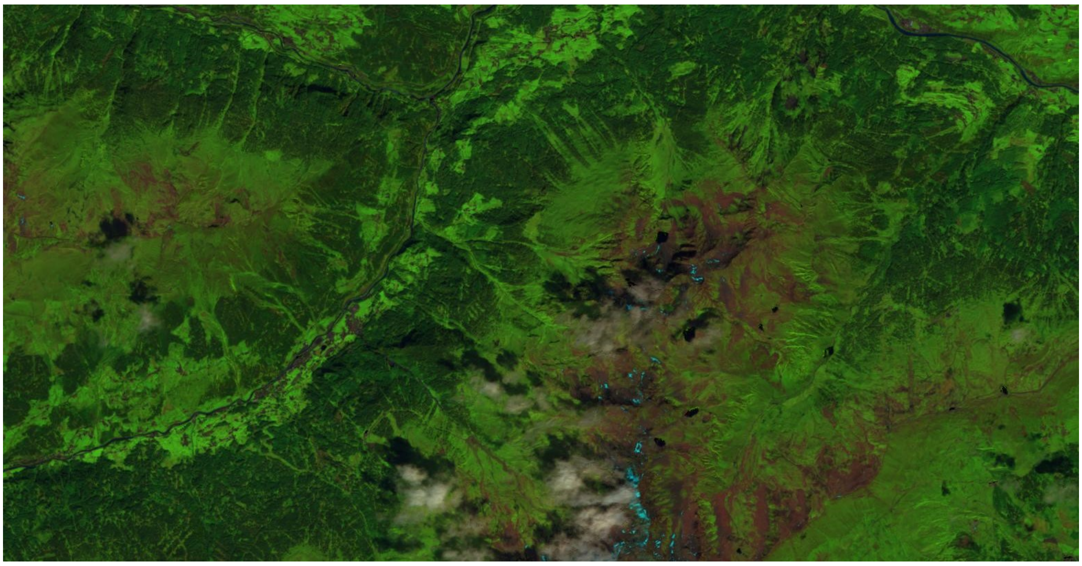
(d)



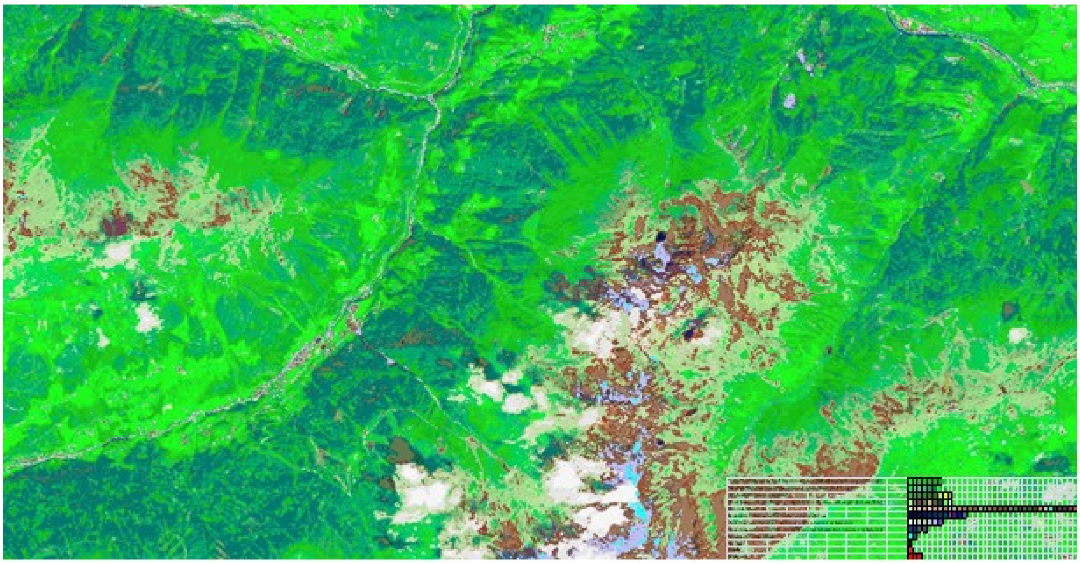
(e)



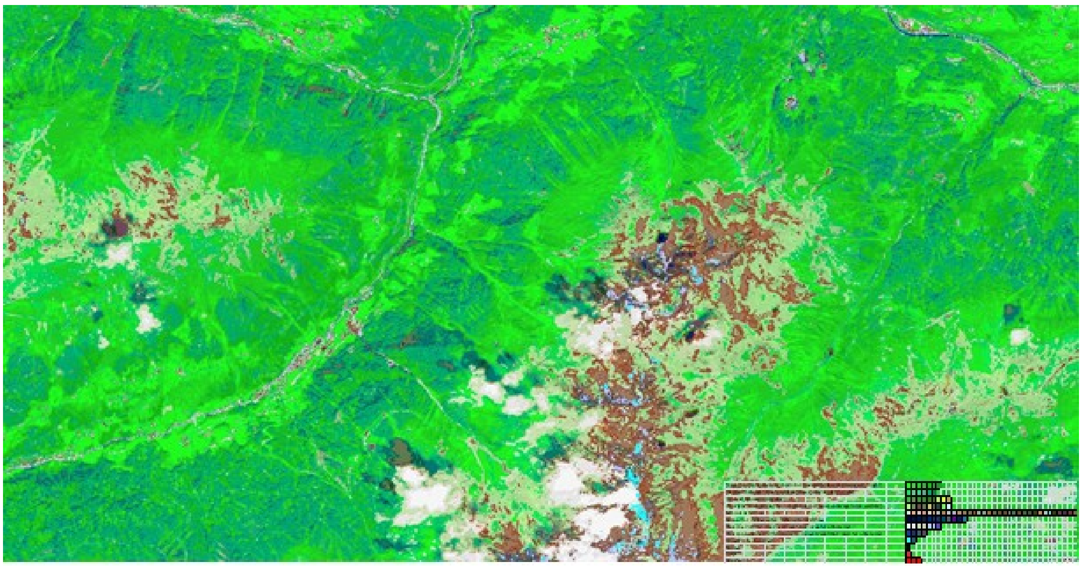
(f)



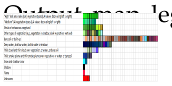
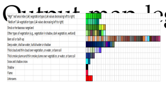
(g)



(h)



(i)

Figure 3. Automated stratified topographic correction (STRATCOR) of TOARF into SURF values, see Figure 4, e.g., refer to [11,80]. (a) Input data set 1: Digital Terrain Model (DTM) of Tirol, Austria. 10 m resolution. (b) Input data set 2: Sentinel-2 Multi-Spectral Instrument (MSI) image of Austria (acquisition date: 2016-08-07), depicted in false colors (R: band MIR, G: band NIR, B: band Blue), 10 m resolution, calibrated into TOARF values. No histogram stretching is applied for visualization purposes. (c) Input data set 3: Satellite Image Automatic Mapper (SIAM) output map in color names [13–15,80,83,84,111–116], automatically generated from the input Sentinel-2 image in TOARF values. Output map legend at coarse granularity, 18 spectral categories, depicted in false colors:  (d) STRATCOR output product. Sentinel-2 MSI image of Austria, depicted in false colors (R: band MIR, G: band NIR, B: band Blue), 10m resolution, radiometrically calibrated into TOARF values and automatically corrected for topographic effects. No histogram stretching is applied for visualization purposes. (e) Zoom-in of the DTM shown in (a). (f) Zoom-in of the input Sentinel-2 image in TOARF values shown in (b). (g) Zoom-in of the STRATCOR output image shown in (d). To be compared with the zoomed-in input image shown in (f). (h) Zoom-in of the SIAM output map in color names, automatically generated from the input Sentinel-2 image in TOARF values shown in (b). Output map legend at coarse granularity, 96 spectral categories, depicted in false colors: . (i) Zoom-in of the SIAM output map in color names, automatically generated from the STRATCOR output image in TOARF values shown in (d). Output map legend at coarse granularity, 96 spectral categories, depicted in false colors: . To be compared with the zoomed-in SIAM map generated from the input image, shown in (h).This is an example of the radiometric correction approach sketched in Figure 12 of the Part 1, where preliminary classification and correction stages alternate in a hierarchical sequence.

II. Required input data set: panchromatic image or MS image provided with no radiometric *Cal* metadata file. In this case, an inherently ill-posed color constancy algorithm must be run for image pre-processing to guarantee input data interoperability (harmonization, reconciliation) across time, space and sensors, see Figure 4. In human vision [11,13,35,47–63], color constancy ensures that the perceived color of objects remains relatively constant under varying illumination conditions, so that they appear identical to a "canonical" (reference) image subject to a "canonical" (known) light source (of controlled quality), e.g., under a white light source [118]. In short, solution of the color constancy problem is the recovery "of an illuminant-independent representation of the reflectance values in a scene" [119]. In practice, color constancy supports image harmonization and interoperability when a radiometric *Cal* metadata file is not available to transform dimensionless DNs into a physical unit of radiometric measure. Computational color constancy is an under-constrained problem in the Hadamard sense [118,120], hence it is difficult to solve. Since it does not have a unique solution, color constancy requires *a priori* knowledge in addition to sensory data to become better conditioned for numerical solution [41,42]. Unfortunately, biophysical mechanisms of color constancy remain largely unknown to date. Hence, to become better posed for numerical solution, inherently ill-posed computational color constancy algorithms cannot be constrained by prior knowledge of color constancy mechanisms in biological vision. For these reasons, a large number of alternative color constancy algorithms have been proposed in the CV literature in the last 30 years [118–120]. Inspired by human vision, a novel self-organizing statistical algorithm for multi-band image color constancy was recently implemented, as reported in [13,121,122].

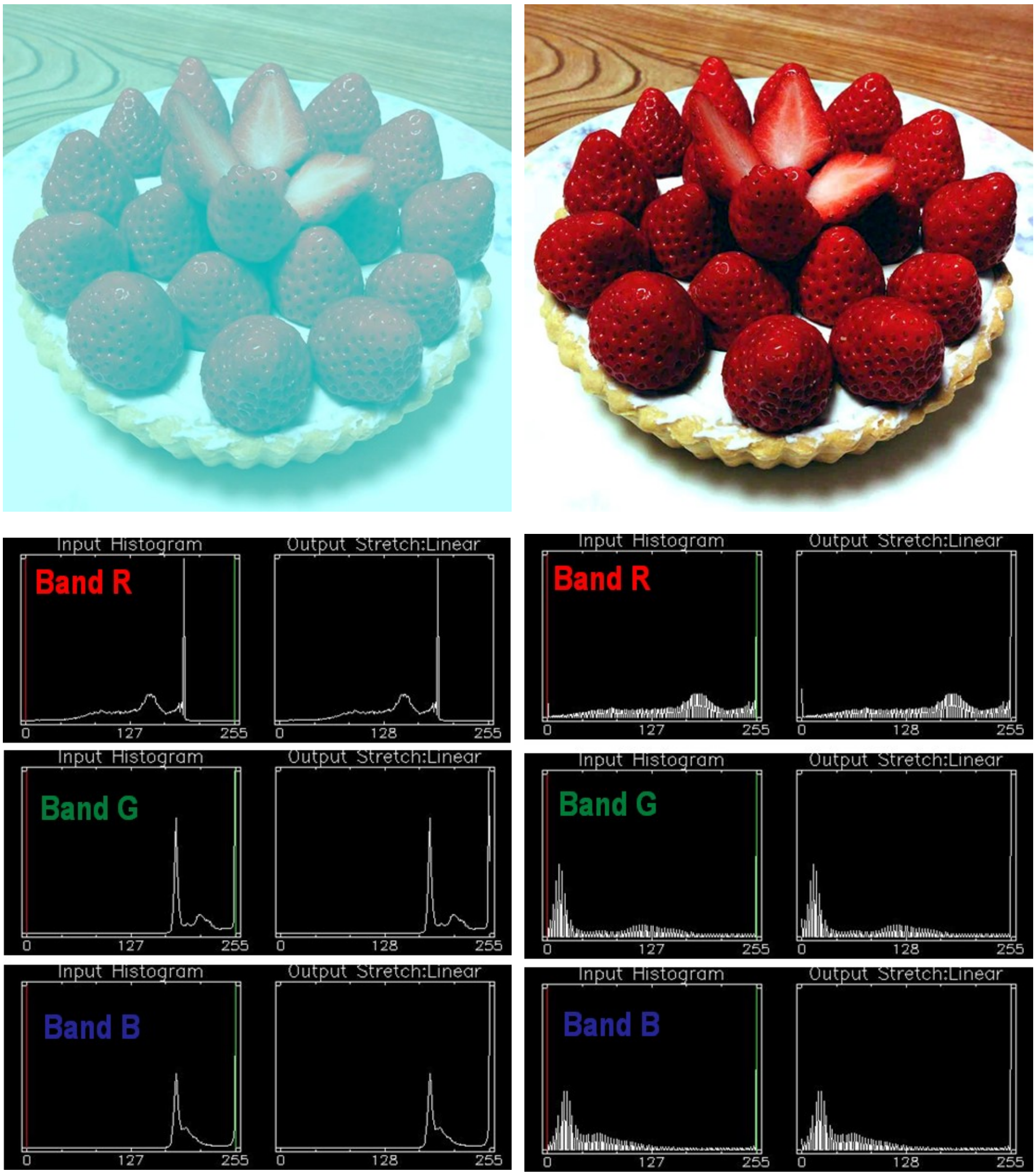
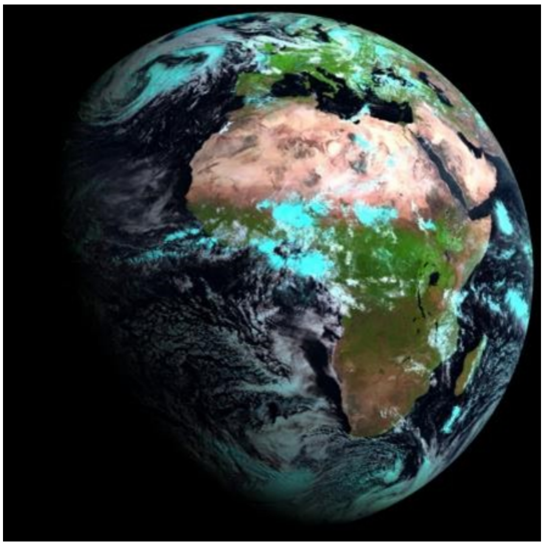


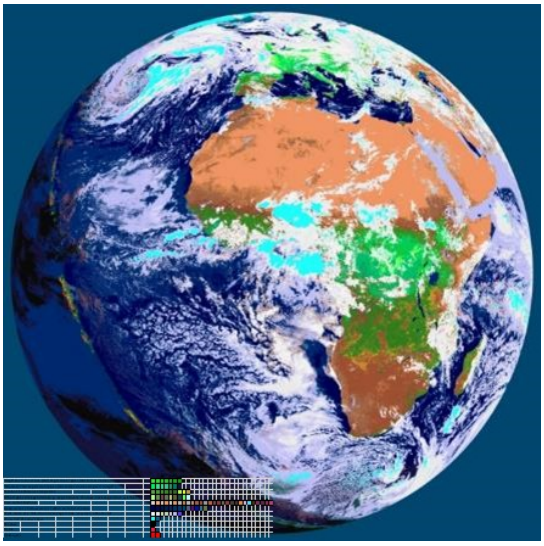
Figure 4. Top left: RGB image, source: Akiyoshi Kitaoka @AkiyoshiKitaoka, web page: http://nymag.com/selectall/2017/02/strawberries-look-red-without-red-pixels-color-constancy.html. Strawberries appear to be reddish, though the pixels are not, refer to the monitor-typical RGB input-output histograms shown at bottom left. No histogram stretching is applied for visualization purposes, see the monitor-typical RGB input-output histograms shown at bottom left. Top right: Output of the self-organizing statistical model-based color constancy algorithm, as reported in [13,121,122], input with the image shown top left. No histogram stretching is applied for visualization purposes, see the monitor-typical RGB input-output histograms shown at bottom right.

(ii) First-stage prior knowledge-based (deductive) color naming.

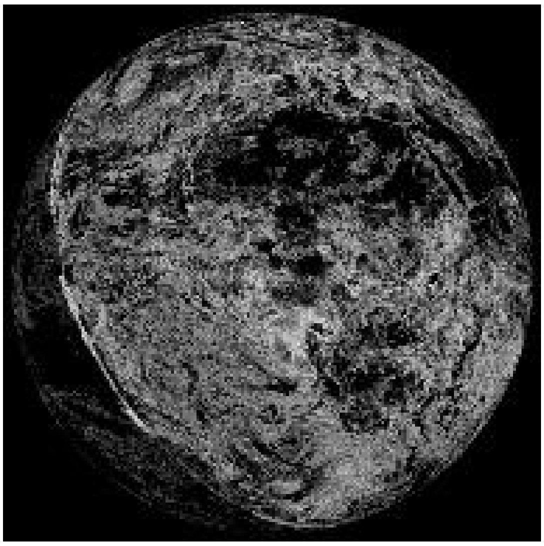
I. Required input data set: MS image radiometrically calibrated into TOARF, SURF or surface albedo values, in agreement with the GEO-CEOS QA4EO *Cal/Val* requirements [2]. Proposed to the RS community in recent years, the Satellite Image Automatic Mapper™ (SIAM™) is a lightweight computer program for automated near real-time MS reflectance space hyperpolyhedralization into MS color names, superpixel detection and vector quantization (VQ) quality assessment [13–15,80,83,84,111–116], see Figure 5, Figure 6 and Figure 7. SIAM claims its scalability to MS imaging sensors featuring different spectral resolution specifications, see Table 1 and Table 2. Moreover, the SIAM's spectral knowkedge-based decision tree for MS reflectance space hyperpolyhedralization outperforms its counterparts in terms of spectral quantization capability at different quantization levels, parameterized by the total number of detected color names at different levels of color granularity, see Table 1 in comparison with Table 1 in the Part 1 for Sen2Cor [16,17], Table 3 in the Part 1 for ATCOR [21,22] and Table 4 in the Part 1 for ATCOR-SPECL [23]. To accomplish a superior scalability to changes in sensor's specifications (e.g., spectral resolution and per-channel curve of sensitivity) and a superior robustness to changes in input data, ranging from TOARF, SURF to surface albedo values, SIAM features a superior degree of redundancy of its multivariate spectral rule set [83,84] in comparison with Sen2Cor, ATCOR and SPECL's. For example, among the spectral knowledge-based decision trees under comparison, the SIAM decision tree is the sole to adopt two different sets of spectral rules to model the multivariate shape and the multivariate intensity information components of a target manifold (hypervolume) of MS signatures, see Figure 18 in the Part 1.



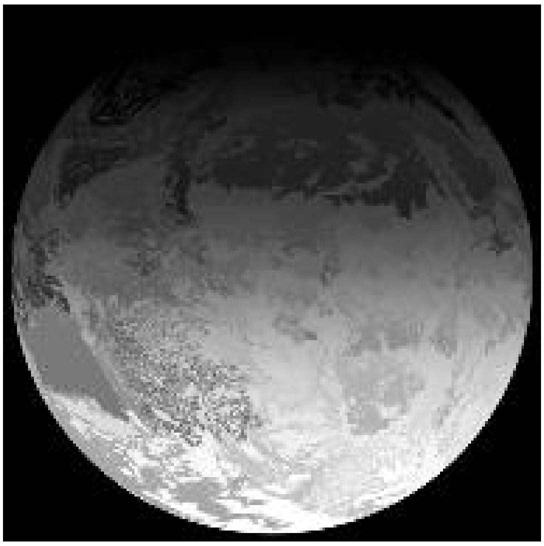
(a)



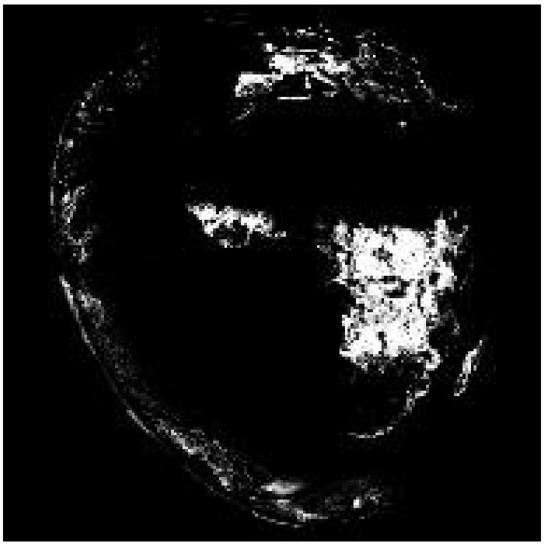
(b)



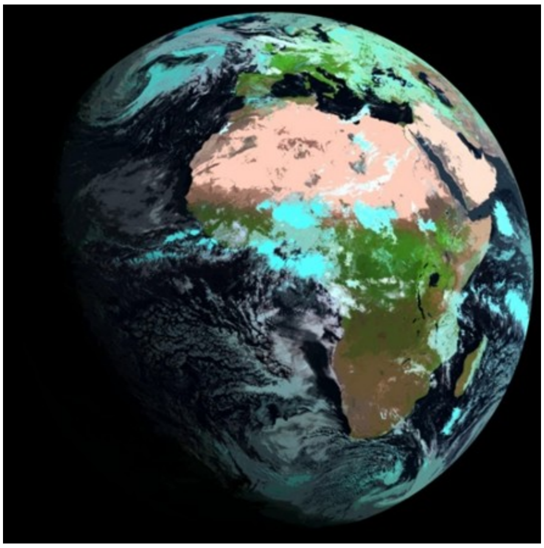
(c)



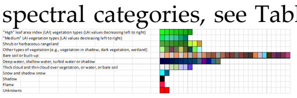
(d)



(e)



(f)

Figure 5. (a) Meteosat Second Generation (MSG) SEVIRI image acquired on 2012-05-30, radiometrically calibrated into TOARF values and depicted in false colors (R: band MIR, G: band NIR, B: band Blue), spatial resolution: 3 km. No histogram stretching is applied for visualization purposes. (b). Advanced Very High Resolution Radiometer (AVHRR)-like SIAM (AV-SIAM™, release 88 version 7) hyperpolyhedralization of the MS reflectance hyperspace and prior color map of the input MS image. The AV-SIAM map legend, consisting of 83 spectral categories, see Table 1, is depicted in pseudocolors. Map legend, similar to Table 2:  (c) To visualize contours of image-segments automatically detected in the multi-level SIAM color map-domain by a deterministic two-pass connected-component multi-level image labeling algorithm [12,123], see Figure 6, an automatic 4- or 8-adjacency cross-aura measure is estimated in linear time, see Figure 7. (d) Segmentation map deterministically generated from the SIAM multi-level output map shown in (b). Each segment is identified by a monotonically increasing (from top to bottom) integer number. (e). Vegetation binary mask, automatically generated from the SIAM multi-level output map shown in (b). On the left, pixel candidates for vegetation belong to spectral category: vegetation in shadow. (f) Piecewise-constant image reconstruction, where each pixel is replaced by the mean reflectance value of the segment that pixel belongs to ("object-mean view"). If the color quantization error (equal to the per-pixel absolute difference between the

input image and the piecewise-constant image reconstruction) is "low", then the quality of the prior knowledge-based SIAM's color space partitioning is "high". No histogram stretching is applied for visualization purposes.

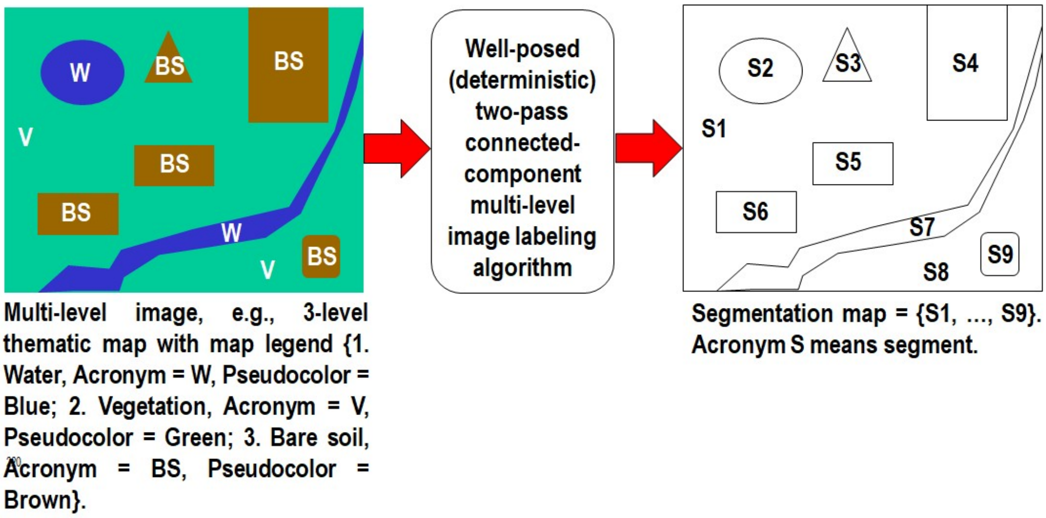


Figure 6. One segmentation map is deterministically generated from one multi-level (e.g., binary) image, such as a thematic map, but the vice versa does not hold, i.e., many multi-level images can generate the same segmentation map. To accomplish the determinist task of segmentation map generation from a multi-level image, the two-pass connected-component multi-level image labeling algorithm [12,123] requires two raster scans of the input data set. In the figure above, as an example, nine image-objects/segments S1 to S9 can be detected in the 3-level thematic map shown at left. Each segment (image-object) consists of a connected set of pixels sharing the same multi-level map label. An image-object is either (0D) pixel, (1) line or (2D) polygon [100]. Each stratum/layer/level consists of one or more segments, e.g., stratum Vegetation (V) consists of two disjoint segments, S1 and S8. In general, a stratum is a multi-part polygon. Hence, in any multi-level (categorical, nominal, qualitative) image domain, three labeled spatial primitives (spatial units) co-exist and are provided with parent-child relationships: (i) pixel with a level-label and a pixel identifier (ID, e.g., the row-column coordinate pair), (ii) segment (either 0D, 1D, or 2D) with a level-specific label and a segment ID, and (iii) stratum (multi-part polygon) with a level-specific 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 [46].

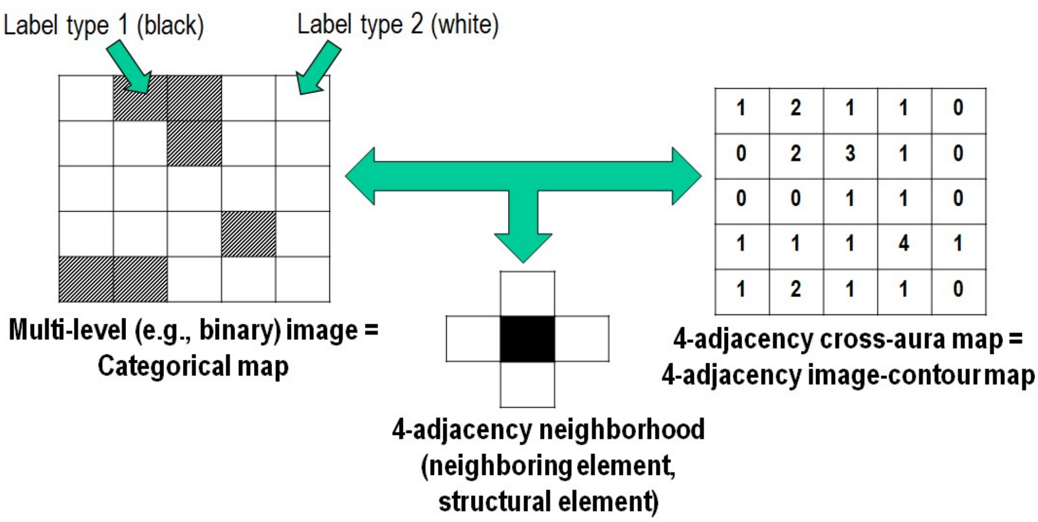
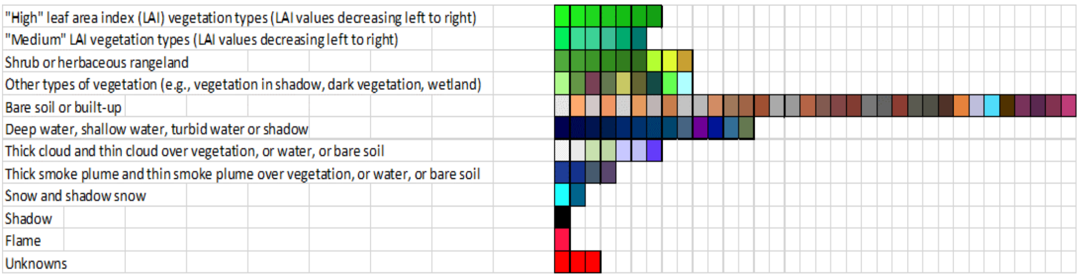


Figure 7. Example of a 4-adjacency cross-aura map, shown at right, generated in linear time from a multi-level (e.g., two-level, binary) image shown at left [11].

Table 1. The SIAM computer program (release 88 version 7) is an EO system of systems scalable to any past, existing or future MS imaging sensor provided with radiometric calibration metadata parameters. 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™) [13–15,80,83,84,111–116].

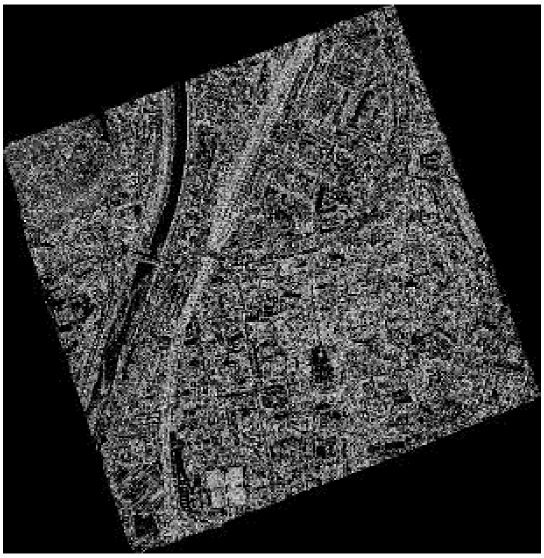
SIAM, r88v7	Input bands	Prior knowledge-based color map legends: Number of output spectral categories = Vocabulary of multi-spectral (MS) color names			
		Fine discretization levels	Intermediate discretization levels	Coarse discretization levels	Inter-sensor discretization levels (*)
L-SIAM	7 – B, G, 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	

Table 2. Legend (vocabulary) of the prior knowledge-based color map generated from a 7-band MS image (consisting of Landsat-like bands B, G, R, NIR, MIR1, MIR2 and TIR), radiometrically calibrated into TOARF, SURF or surface albedo values, by the Landsat-like SIAM (L-SIAM™, release 88 version 7) implementation, also refer to Table 1. For the sake of representation compactness, pseudo-colors of the 96 spectral categories are gathered along the same raw if they share the same parent spectral category in the decision tree, e.g., "strong" vegetation, equivalent to a spectral end-member. The pseudo-color of a spectral category (color name) is chosen as to mimic natural colors of pixels belonging to that spectral category.



II. Required input data set: RGB image, either true- or false-color, not provided with radiometric calibration metadata parameters, but submitted to a self-organizing color constancy algorithm for harmonization purposes. The RGB Image Automatic Mapper™ (RGBIAM™) is a lightweight computer program capable of color naming, superpixel detection and VQ quality assessment in non-calibrated RGB imagery, either true- or false-color, submitted to color constancy as a mandatory non-calibrated RGB image pre-processing stage for image harmonization across space, time and sensors [13,121,122]. RGBIAM comprises a novel expert system (*a priori* spectral knowledge-based decision tree) capable of partitioning a three-band RGB data cube, either true- or false-color, into a pre-defined vocabulary of RGB color names (equivalent to a discrete and finite set of mutually exclusive and totally exhaustive polyhedra, neither necessarily convex nor connected) at two different quantization levels of color granularity, specifically 50 and 11 color names, see Figure 8 and Table 3 [13,121,122].

(a) (b)



(c) (d)

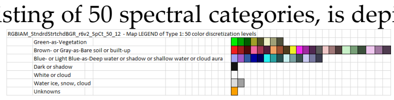
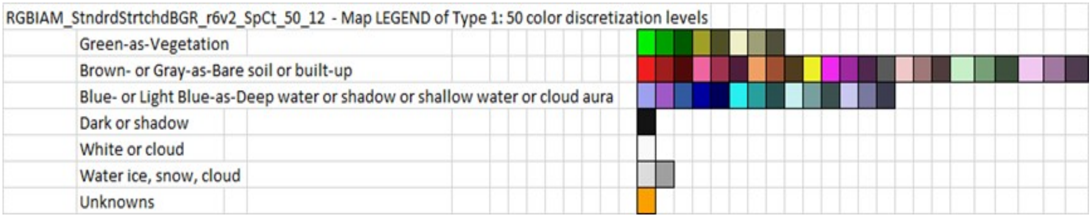
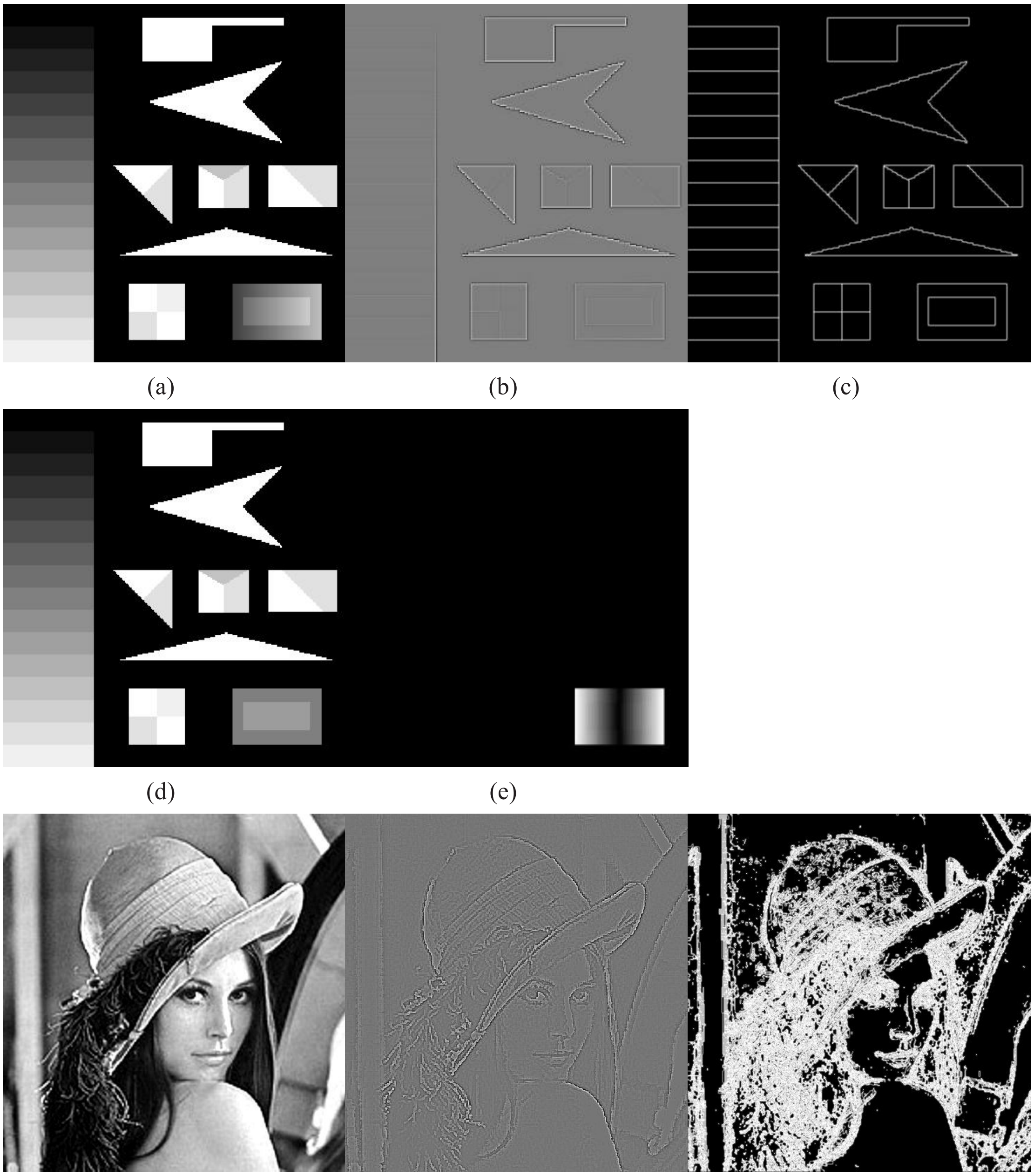
Figure 8. (a) Airborne 10 cm resolution true-color RGB orthophoto of Trento, Italy, 4017 x 4096 pixels in size x 3 bands, acquired in 2014 and provided with no radiometric calibration metadata file. No histogram stretching is applied for visualization purposes. (b) Same RGB orthophoto subject to self-organizing statistical color constancy. (c) RGBIAM polyhedralization of the RGB color space and prior color map of the RGB image subject to color constancy. The RGBIAM map legend, consisting of 50 spectral categories, is depicted in pseudocolors. Map legend, shown in Table 3: . Input parameters: none. Processing time (one-pass, IDL implementation) = 2 min. (d) To visualize contours of image-segments automatically detected in the multi-level RGBIAM color map-domain by a deterministic two-pass connected-component multi-level image labeling algorithm [12,123], see Figure 6, an automatic 4- and/or 8-adjacency cross-aura measure is estimated in linear time, see Figure 7.

Table 3. Legend (vocabulary) of the prior knowledge-based color map generated from a 3-band RGB image, in either true- or false-color and submitted to a color constancy pre-processing stage, by the RGBIAM (release 6 version 2) implementation at fine color quantization with 50 color names (versus 11 basic color names at coarse color space partitioning). For the sake of representation compactness, pseudo-colors of the 50 spectral categories are gathered along the same raw if they share the same parent spectral category in the decision tree. The pseudo-color of a spectral category (color name) is chosen as to mimic natural colors of pixels belonging to that spectral category.



(iii) Second stage for image-contour detection and image-segmentation (raw primal sketch) [11,13]. Based on an original physical model-based multi-scale multi-orientation 2D wavelet-based spatial filter bank [11], equivalent to a prior knowledge-based (deductive) CNN alternative to inductive learning-from-data DCNNs currently dominating the CV literature and whose architecture is based on heuristics, the proposed physical model-based CNN is consistent with constraint 'Human vision → CV ⊃ EO-IU', see Figure 2 in the Part 1. In particular, a CV ⊃ EO-IU subsystem is tested on complex EO spaceborne/airborne images if and only if it performs in agreement with human visual perception on test cases at increasing levels of signal complexity, e.g., 1D synthetic signal, (2D) synthetic image, natural panchromatic imagery and natural color imagery whose "ground truth" is intuitive to understand, etc. Perceptual criteria to comply with are: (I) the perceptual true fact that human panchromatic and chromatic vision mechanisms are nearly as effective in scene-from-image reconstruction and understanding (refer to Section 2 in the Part 1), and (II) the Mach bands visual illusion [11,124], see Figure 9 in the Part 1. In the words of Serre, "there is growing consensus that optical illusions are not a bug but a feature. I think they are a feature. They may represent edge cases for our visual system, but our vision is so powerful in day-to-day life and in recognizing objects" [56,57]. Traditional DCNNs only include feedforward connections between layers, not Serre's innovative feedback connections between neurons within a layer, which were found to be necessary and sufficient to model contextual optical illusions in human vision [56,57]. The proposed physical model-based multi-scale multi-orientation 2D spatial filter bank (convolutional neural network, CNN) design and implementation complies with the two aforementioned perceptual requirements. To reduce computation time, each 2D wavelet-based spatial filter, 2D_f(·), assumed to be W × W pixels in size, is implemented as a 2D separable filter, such that 2D_f(x,y) = 1D_g(x) × 1D_g(y), where the 1D wavelet-based spatial filter, 1D_g(·), is W pixels in size [53]. Hence, computational complexity of separable per-filter image convolution is 2W × image size in pixels, rather than W × W × image size in pixels. On theory, four filter orientations (0°, 45°, 90° and 135°) and four (up to seven) dyadic spatial scales, s = 0, …3, are employed in agreement with biological evidence [54, 58–63], where filter size $W_0 = 2^0 \times 3 = 3$ pixels, $W_1 = 2^1 \times 3 \approx 7$ pixels, $W_2 = 2^2 \times 3 \approx 13$ pixels and $W_3 = 2^3 \times 3 \approx 25$ pixels. In practice, the same W_0 filter is applied hierarchically upon the raw image at scale 0 and on three low-pass filtered images, downscaled by a factor of 2 at scales 1 to 3 (eventually, down to scale 7) [53]. The implemented physical model-based CNN is capable of: (a) near-orthogonal image analysis/decomposition and lossless image synthesis/reconstruction, in analogy with the well-known multi-scale Gaussian and Laplacian pyramid proposed in [53], (b) automated (requiring no human–machine interaction) zero-crossing (ZX) image-contour detection [11], in line with a biologically plausible [11,35,47–63] CV system proposed by Marr [13], and (c) automated ZX image-segment (blob, closed-contour) detection [11], never accomplished by Marr in his seminal work [13], see Figure 9. Such a deductive CNN implementation is alternative to inductive learning-from-data algorithms for image-contour detection and image segmentation, either 2D image analysis approaches, such as DCNNs [65,68–71], or 1D image analysis approaches, e.g., [125–127], which are inherently semi-automatic and site-specific [128].



(a) (b) (c)

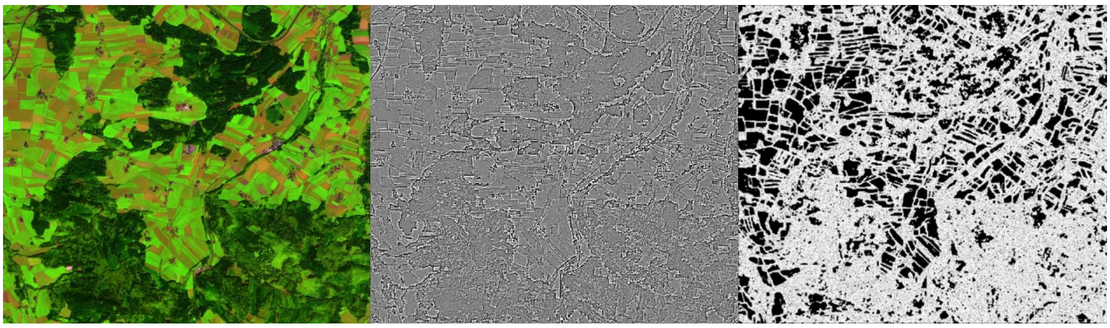
(d) (e)

(f) (g) (h)

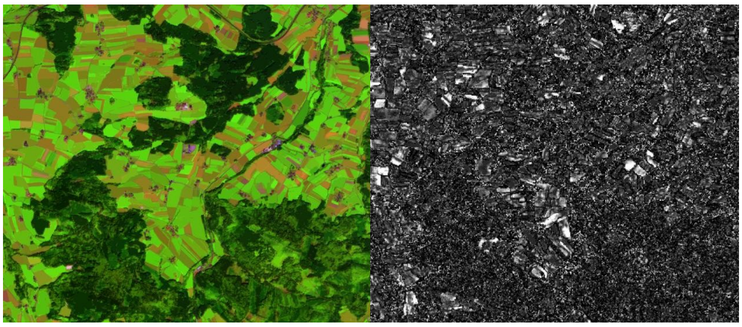
(i) (l)

(m) (m) (o)

(p) (q)



(r) (s) (t)



(u) (v)

Figure 9. To comply with constraint 'Human vision → CV ⊃ EO-IU', see Figure 2 in the Part 1, an EO-IU subsystem for image-contour detection and image segmentation is tested on complex EO spaceborne/airborne images *if and only if* it performs in agreement with human visual perception [56,57], such as: (1) the Mach bands visual illusion [124] and (2) the perceptual true fact that human panchromatic and chromatic vision mechanisms are nearly as effective in scene-from-image reconstruction and understanding, starting from simpler test of increasing signal complexity, e.g., 1D synthetic signal, 2D synthetic image, natural panchromatic and natural chromatic images of intuitive "ground truth", etc. (a) SUSAN synthetic panchromatic image, byte coded in range {0, 255}. Step edges and ramp edges at known locations (the latter forming the two inner rectangles visible at the bottom right corner) form angles from acute to obtuse. According to human vision, 31 image-segments can be detected as reference "ground-truth". (b) Sum (synthesis) of the wavelet-based near-orthogonal multi-scale multi-orientation image decomposition. Filter value sum in range from -255.0 to +255.0. (c) Automated (requiring no human–machine interaction) image segmentation into zero-crossing (ZX) segments generated from ZX pixels detected by a multi-scale multi-orientation spatial filter bank, equivalent to a prior knowledge-based CNN, different from Marr's single-scale isotropic ZX pixel detection [13]. Exactly 31 image-segments are detected with 100% contour accuracy. Segment contours depicted with 8-adjacency cross-aura values in range {0, 8}, see Figure 7. (d) Image-object mean view = object-wise constant input image reconstruction. (e) Object-wise constant input image reconstruction compared with the input image, per-pixel root mean square error (RMSE) in range 0.0–255.0. (f) Natural panchromatic image of Lenna. (g) Same as (b). (h) Same as (c), there is no CV system's free-parameter to be user-defined. (i) Same as (d). (l) Same as (e). (m) Natural RGB-color image of Lenna. (n) Same as (b). (o) Same as (c), there is no CV system's free- parameter to be user-defined. (p) Same as (d). (q) Same as (e). (r) Zoom-in of a Sentinel-2A MSI Level-1C image of the Earth surface south of the city of Salzburg, Austria. Acquired on 2015-09-11. Spatial resolution: 10 m. Radiometrically calibrated into top-of-atmosphere reflectance (TOARF) values in range {0, 255}, it is depicted as a false color RGB image, where: R = Medium InfraRed (MIR) = Band 11, G = Near IR (NIR) = Band 8, B = Blue = Band 2. Standard ENVI histogram stretching applied for visualization purposes. (s) Same as (b). (t) Same as (c), there is no CV system's free-parameter to be user-defined. (u) Same as (d). (v) Same as (e).

(iv) Third stage for planar shape indexing. An original minimally dependent and maximally informative (mDMI) set of planar shape (geometric) indexes was conceived and implemented as described in [11,12,92]. The proposed mDMI set of geometric functions comprises scale-invariant roundness (compactness and no holiness) in range 0.0–1.0, elongatedness (and no holiness) ≥ 1, multi-scale straightness of boundaries in range 0.0–1.0, simple connectivity (no holiness) in range 0.0–1.0, rectangularity (and no holiness) in range 0.0–1.0 and convexity (and no holiness) in range 0.0–1.0, to be estimated per image-object in addition to size and orientation, see Figure 10.

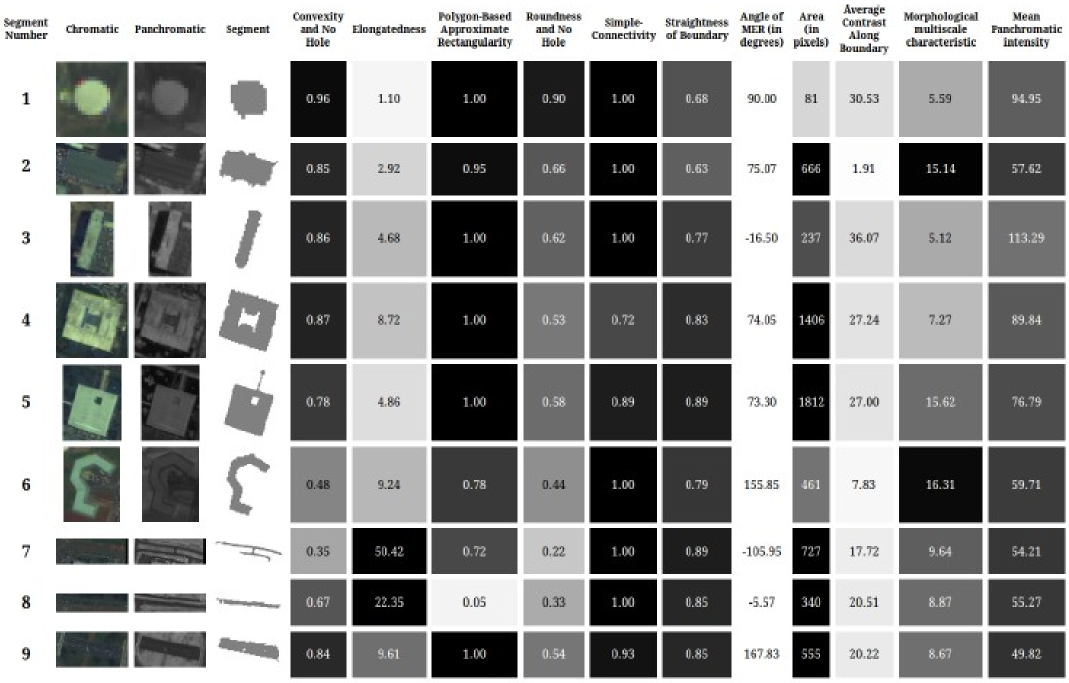


Figure 10. Screenshot of a graphic user interface (GUI) specifically developed to show a human expert values of the proposed minimally dependent and maximally informative (mDMI) set of geometric attributes. In this GUI, darker cells correspond to: (i) higher values of geometric attributes and (ii) lower values of photometric attributes, like the panchromatic mean intensity shown at the rightest column. In this figure, for reasons of readability only nine segments are shown simultaneously for comparison. Detected in the spaceborne very high resolution (VHR) test image of an urban area, segments 1 through 6 correspond to buildings or parts of buildings while segments 7 through 9 belong to roads. These two families of segments appear easy to discriminate based on different combinations of ranges of change of their geometric attributes.

(v) Fourth stage for texture segmentation (full primal sketch), synonym for perceptual spatial grouping of texture elements, texels [11,12,93–95] or tokens [13]. Based on a multi-scale texture binary profile accomplished in linear time complexity via multi-scale window-based analysis of the spatial distribution of binary image-contours [11]. To date, automated image-texture segmentation (perceptual grouping of texels) is an open problem [13,25,26,54,129].

(vi) Fifth stage for general-purpose, user- and application-independent ESA EO Level 2 product generation (refer to Section 1 in the Part 1), based on a 2D convergence-of-visual-evidence approach, according to Equation (2) [11], where the proposed ESA EO Level 2 SCM product taxonomy consists of a standard 3-level 8-class FAO LCCS-DP legend augmented with quality layers cloud and cloud–shadow, see Table 2 and Figure 1 in the Part 1.

(vii) Sixth stage for user- and application-dependent FAO LCCS-MHP classification (see Figure 1 in the Part 1), based on a 2D convergence-of-visual-evidence approach, according to Equation (2) [11].

At the four levels of understanding known as information/knowledge representation, system architecture, algorithm and implementation (refer to Section 1), the AutoCloud+ CV system can be considered a specific instantiation of the aforementioned six-stage hybrid feedback CV system design, shown in Figure 2. In more detail, the AutoCloud+ flow chart, algorithm and implementation are instantiated as follows [11,76,77], see Figure 11.

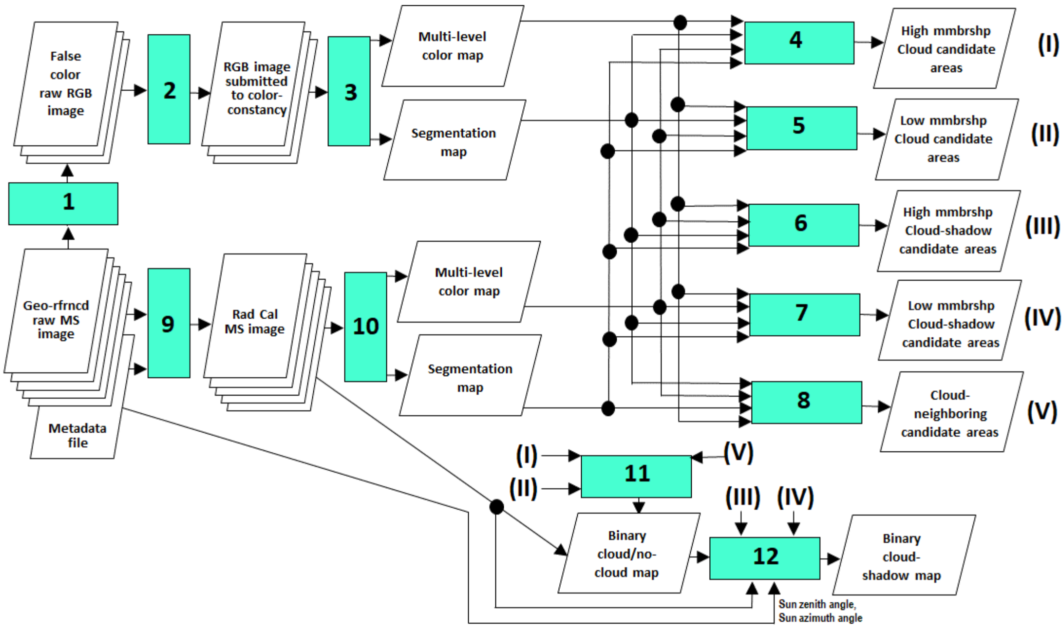


Figure 11. AutoCloud+ hybrid (combined physical model-based and statistical model-based) 2D image analysis system design (architecture) for spatial context-sensitive and spatial topology-preserving cloud/cloud–shadow detection. (1) True- or false-color RGB channel selection. (2) Statistical self-organizing color constancy algorithm. (3) RGBIAM lightweight computer program for RGB color space polyhedralization into color names, superpixel detection and vector quantization (VQ) quality assessment. (4) Candidate cloud areas, based on convergence-of evidence. (5) Candidate no-cloud areas, based on convergence-of evidence. (6) Candidate cloud–shadow areas, based on convergence-of evidence. (7) Candidate no-cloud–shadow areas, based on convergence-of evidence. (8) Candidate cloud neighboring areas, based on convergence-of evidence. (9) Radiometric calibration of digital numbers (DNs) into TOARF, SURF or surface albedo values, in compliance with the GEO-CEOS QA4EO *Cal/Val* requirements [2,130]. (10) SIAM lightweight computer program for multi-spectral (MS) reflectance space hyperpolyhedralization into color names, superpixel detection and VQ quality assessment. (11) Spatial cloud modeler: Clouds detected from candidate cloud and cloud neighboring areas. (12) Spatial bidirectional cloud and cloud–shadow modeler: Physical model-based cloud shadow detection, moving from cloud to cloud–shadow candidates and vice versa, for mutual reinforcement learning.

First, the two SIAM and RGBIAM lightweight computer programs for color space discretization into color names are combined to pursue convergence of photometric (colorimetric) evidence, whether or not the color space is radiometrically calibrated [130]. For example, when the input MS image is radiometrically calibrated into TOARF, SURF or surface albedo values in compliance with the GEO-CEOS QA4EO *Cal/Val* requirements [2], then SIAM and RGBIAM are run in parallel upon, respectively, the original input MS image (calibrated) and an RGB false-color selection of the input (calibrated) image pre-processed by the self-organizing algorithm for color constancy (similar to image enhancement by histogram stretching). The two output multi-level maps in color names generated as output by SIAM and RGBIAM can be deterministically segmented in linear time complexity with image size into labeled image-objects (super-pixels, connected sets of pixels featuring the same color label) by a well-known deterministic (well-posed) two-pass connected-component multi-level image labeling algorithm [12,123], see Figure 6. In the SIAM and RGBIAM output segmentation map, each connected-component is labeled with one segment identifier (ID) together with one semi-symbolic color name. For example, an image-object, either (0D) pixel, (1) line or (2D) polygon [100], is provided with a segment ID, a MS color name 'green-as-Vegetation' by SIAM and an RGB color name 'Green' by RGBIAM.

Second, candidate cloud-objects and candidate cloud–shadow-objects are detected in the image-domain based on a convergence of shape and size properties with SIAM and RGBIAM color names. Stratified (class-conditioned) by color names, a driven-by-knowledge shape-preserving dilation process is run to fill small spatial gaps within cloud/cloud–shadow candidate areas (e.g., due to self-occlusion and shadow-casting phenomena occurring in clouds, where solar luminance decreases) and along boundaries of candidate cloud areas.

Third, the cloud candidate image-objects can grow in the image-domain to include spatially adjacent image-objects whose color, shape and size properties are consistent with the hypothesis of belonging to class crown-of-cloud.

Finally, a bidirectional physical knowledge-based data modeler is run for spatial reasoning (see software blocks identified with numbers 11 and 12 in Figure 11). Making use of OBIA concepts for spatial information modeling [25], it is capable of cloud/cloud–shadow image-object pair spatial matching in shape while accounting for the sun position, to reduce cloud and cloud–shadow false positives and false negatives. In the quest for a unidirectional cloud-to-cloud–shadow relationship, candidate cloud–shadows are searched for in the (2D) image-domain to be matched in shape by each individual candidate cloud-object adopted as starting position. This spatial search moves from the candidate cloud-object of interest, in the direction of the sun azimuth angle (known from the input EO image metadata file) with orientation away from the sun, for a spatial length estimated as a dependent variable of the cloud height, which is the sole unknown independent physical variable to cope with, estimated according to an *a priori* physical knowledge-based model of real-world cloud heights, in line with Sen2Cor [17]. The dual quest for object-pair matching in shape deals with the cloud–shadow-to-cloud unidirectional relationship, which is usually neglected in other cloud/cloud–shadow detectors, such as Sen2Cor [17] and MAJA [18]. A simplified flow chart of the hybrid AutoCloud+ software toolbox for automated cloud/cloud–shadow detection is shown in Figure 12.

At the implementation level of system understanding, it is worth mentioning that candidate cloud/cloud–shadow image-object pairs are matched in shape according to the mDMI set of planar shape functions implemented in [11,12,92]. The AutoCloud+ software blocks identified as data processing units 4 to 8, 11 and 12 in Figure 11 were implemented within the Trimble's eCognition Developer commercial software environment for fast prototyping OBIA solutions [131].

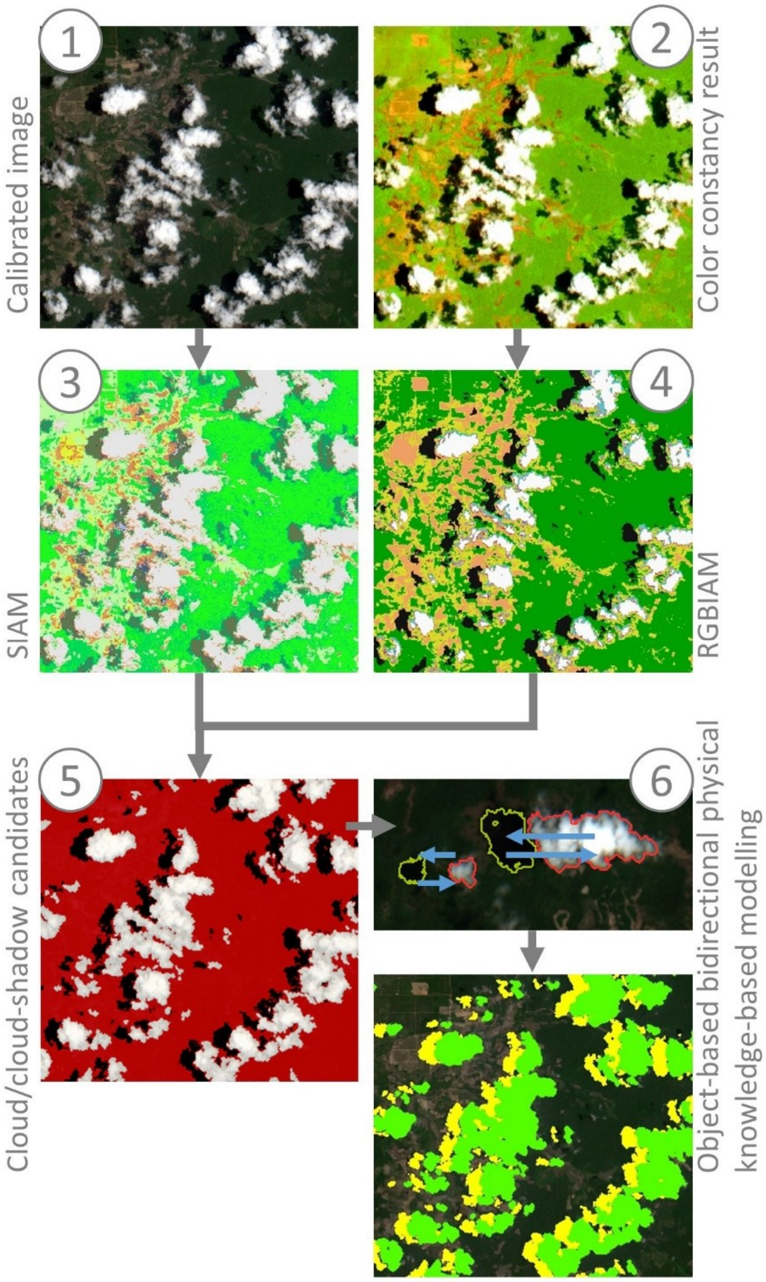


Figure 12. Simplified flow chart of the hybrid AutoCloud+ software toolbox for automated cloud/cloud–shadow detection, see Figure 11. This intuitive workflow is depicted as a sequence of two input data sets (1) and (2), generated from the same input MS image, one submitted to radiometric *Cal* (when radiometric *Cal* metadata parameters are available) and the other submitted to statistical color constancy, followed by four intermediate information products (3) to (6), and one final output product, specifically, a three-level output map with semantic layers cloud/cloud–shadow/others (rest of the world, depicted as the input image), shown at bottom right.

4. Materials

For testing purposes, the novel AutoCloud+ algorithm for joint cloud and cloud–shadow detection in single-date MS imagery was compared in quantitative terms of outcome and process quality, in compliance with the mDMI set of EO outcome and process (OP) quantitative quality indicators (OP-Q²Is) proposed in Section 2 of the Part 1 [11], with alternative standard CV software solutions available open source or free of cost, specifically, the single-date ESA Sen2Cor and multi-temporal MAJA software toolboxes.

As proof of concept, AutoCloud+, Sen2Cor and MAJA (refer to Table 5 in the Part 1) were compared upon two test images, selected to be representative of the complexity of the cloud/cloud–shadow detection problem in real-world situations.

The first test image was selected as a Sentinel-2 A (S2A) Multi-Spectral Instrument (MSI) Level 1C image radiometrically calibrated into TOARF values, depicting an Earth-surface area located in Cambodia (Product ID: S2A_MSIL1C_20170421T031541_N0204_R118_T48PWV_20170421T033212), see Figure 13, for which Sen2Cor and MAJA results were available for download from the Copernicus Open Access Hub [132] and from the CNES website [133] respectively. This first test image shows a large variety of cloud patterns, varying in shape, size and cloud height, together with typical critical elements in cloud and cloud–shadow detection, such as cloud–shadows projected over water and vegetated surface types, in addition to occluded cloud and cloud–shadow phenomena.

A second test image was identified in a Sentinel-2 B (S2B) Level 1C image radiometrically calibrated into TOARF values, depicting an Earth-surface area located in the Alpine area between Austria and Germany (Product ID: S2B_MSIL1C_20180616T102019_N0206_R065_T32TPT_20180616T154713), see Figure 13. In this second test image, potential critical elements in cloud/cloud–shadow detection are the presence of snow/ice patterns in high-elevation areas, eligible for confusion with ice clouds, of clouds located next to visible snow/ice surface types and of Earth surfaces affected by shadows casted by mountains rather than clouds. For this second test image, no MAJA mapping result was available for download from the aforementioned CNES website.

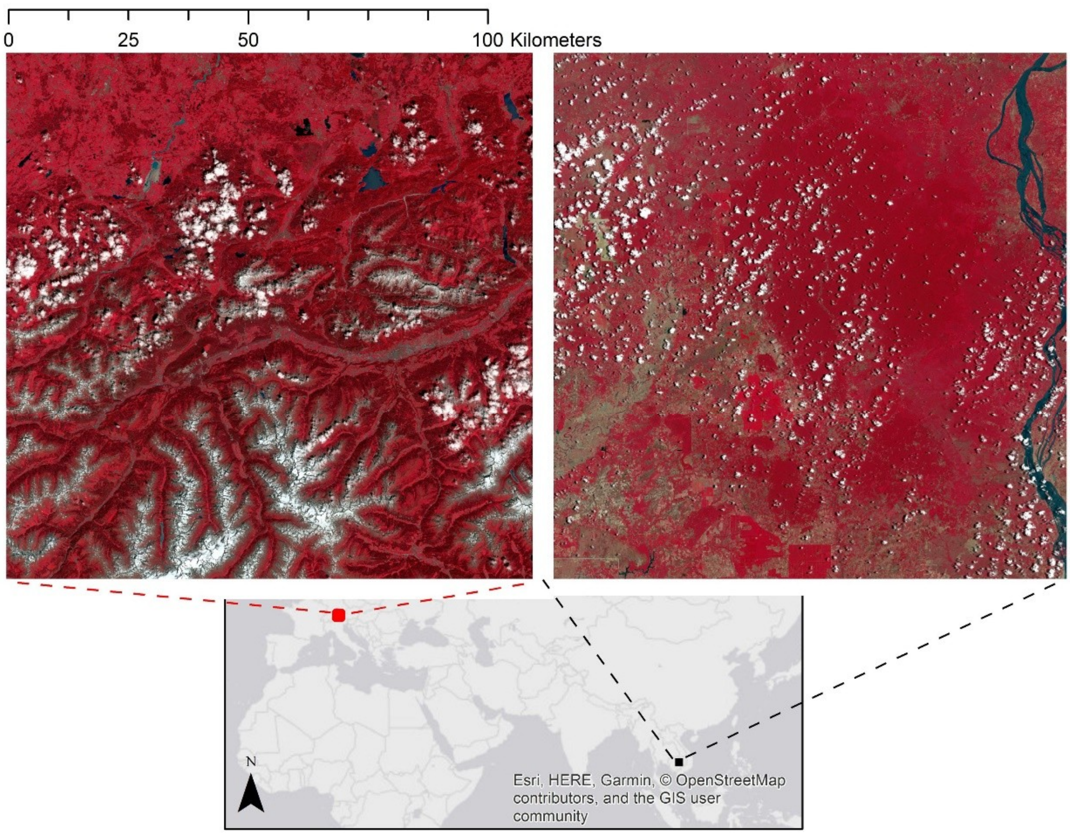


Figure 13. Overview of the two study areas adopted for testing (Left: Cambodia test site, Right: Alpine test site located between Austria and Germany). Test EO images of these two study areas are expected to be representative of the complexity of the cloud/cloud–shadow detection problem in real-world situations. At left, second test image, identified as a Sentinel-2 B Multi-Spectral Instrument (MSI) Level 1C image, radiometrically calibrated into TOARF values, depicted in false-colors: monitor-typical RGB channels are selected as R = Near InfraRed (NIR)

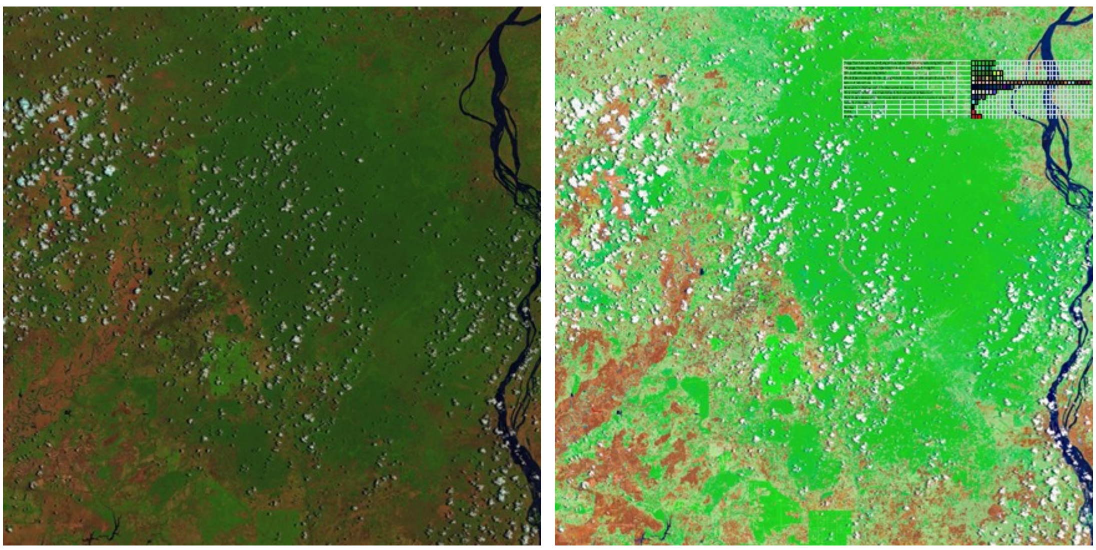
channel, G = Visible Red channel, B = Visible Green channel. Histogram stretching is applied for visualization purposes. Product ID: S2B_MSIL1C_20180616T102019_N0206_R065_T32TPT_20180616T154713. At right, first test image, identified as Sentinel-2 B MSI Level 1C image, radiometrically calibrated into TOARF values, and depicted in false-colors as the second test image shown at left. Product ID: S2A_MSIL1C_20170421T031541_N0204_R118_T48PWV_20170421T033212.

5. Results

Alternative cloud/cloud–shadow output maps generated by the single-date AutoCloud+, single-date Sen2Cor and multi-temporal MAJA software toolboxes were compared in quantitative terms of outcome (product) quality defined in the mDMI set of EO OP-Q²Is proposed in Section 2 of the Part 1. In addition, the AutoCloud+ CV subsystem was assessed in quantitative terms of process quality, in agreement with the mDMI set of OP-Q²Is proposed in Section 2 of the Part 1. Noteworthy, this quality assurance strategy goes far beyond the traditional quantitative assessment policy of EO-IU systems presented in the RS literature, almost exclusively limited to mapping accuracy.

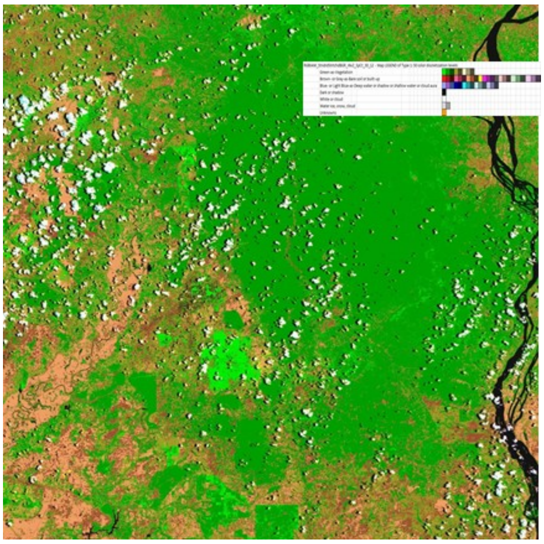
With regard to process quality indicators defined in the mDMI set of OP-Q²Is proposed in Section 2 of the Part 1, such as degree of automation and computation time, AutoCloud+ ran automatically (without human–machine interaction) upon the two test images, where Step 1 to Step 5 in Figure 12 were computed in near real-time, specifically, in linear time complexity with image size, equal to around 2 minutes per Sentinel-2 image in a standard laptop computer. The bidirectional object-based cloud/cloud–shadow spatial modeler, shown as Step 6 in Figure 12 and prototyped in the Trimble eCognition software, depends on the number of cloud and cloud–shadow candidate image-objects detected in the image-domain at Step 5. For each of the two Sentinel-2 images selected for testing, the eCognition software prototype at Step 6 ran in a reasonable time frame, around 5 minutes per input image.

With regard to outcome quality indicators defined in the mDMI set of OP-Q²Is proposed in Section 2 of the Part 1, enhanced input images, either radiometrically calibrated into TOARF values or submitted to color constancy, and intermediate SIAM's and RGBIAM's output maps generated from the first Sentinel-2 test image are shown in Figure 14. Final cloud/cloud–shadow output maps generated by the single-date AutoCloud+, single-date Sen2Cor and multi-temporal MAJA computer programs are shown in Figure 15, zoomed-in in Figure 16 and extra zoomed-in in Figure 17. Table 4 reports on the total area of clouds and cloud–shadows detected by the three algorithms of interest in the first test image.

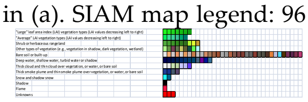
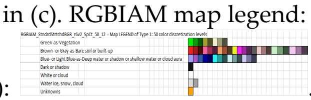


(a) (b)

(c)



(d)

Figure 14. (a) First test image of a Cambodia site. Sentinel-2 A Multi-Spectral Instrument (MSI) Level 1C image of Cambodia, radiometrically calibrated into TOARF values. Acquisition date: 2018-06-16. Depicted in false-colors: monitor-typical RGB channels are selected as R = Medium InfraRed (MIR) channel, G = Near InfraRed (NIR) channel, B = Visible Blue channel. No histogram stretching is employed for visualization purposes. (b) SIAM map of the test S2 image shown in (a). SIAM map legend: 96 color names, depicted in pseudocolors as follows (see Table 2):  (c) First test image, shown in (a), submitted to a self-organizing color constancy algorithm. Depicted in false-colors: monitor-typical RGB channels are selected as R = Medium InfraRed (MIR) channel, G = Near InfraRed (NIR) channel, B = Visible Blue channel. No histogram stretching is employed for visualization purposes. (d) RGBIAM map of the test S2 image shown in (c). RGBIAM map legend: 50 color names, depicted in pseudocolors as follows (see Table 3): .

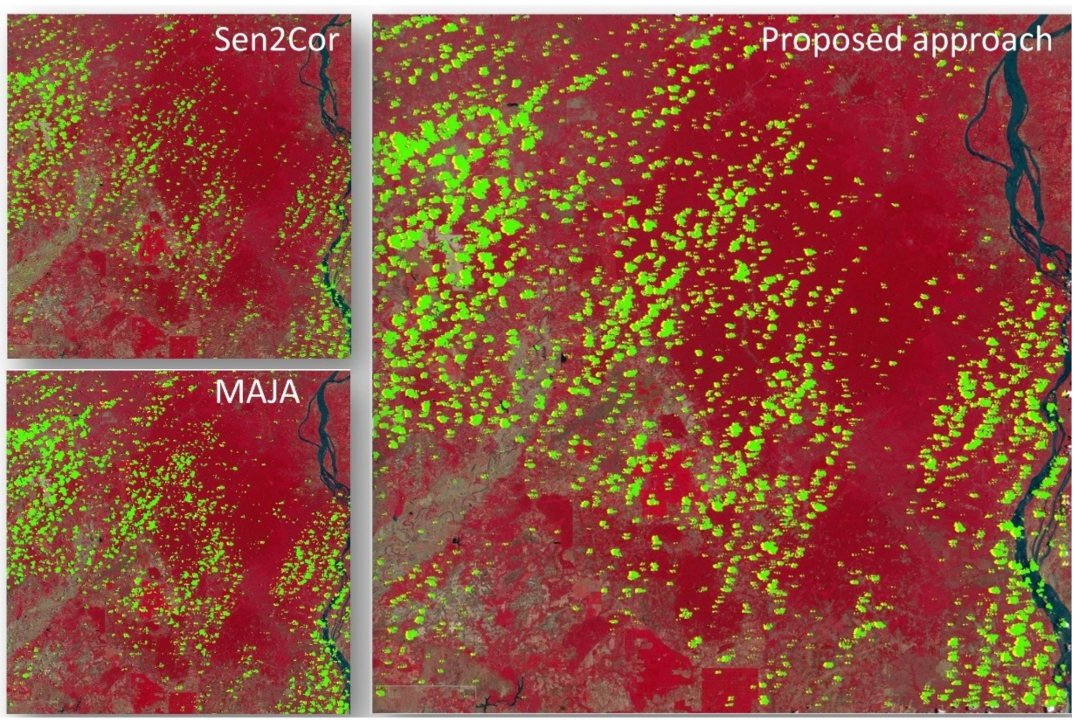


Figure 15. First test image of a Cambodia site. Final 3-level cloud/cloud–shadow/others maps generated by the three algorithms under comparison, specifically, single-date AutoCloud+, single-date Sen2Cor and multi-date MAJA, where class "others" is overlaid with the input Sentinel-2 A Multi-Spectral Instrument (MSI) Level 1C image, radiometrically calibrated into TOARF values and depicted in false-colors: monitor-typical RGB channels are selected as R = Near InfraRed (NIR) channel, G = Visible Red channel, B = Visible Green channel. Histogram stretching is applied for visualization purposes. Output class cloud is shown in a green pseudocolor, class cloud–shadow in a yellow pseudocolor.

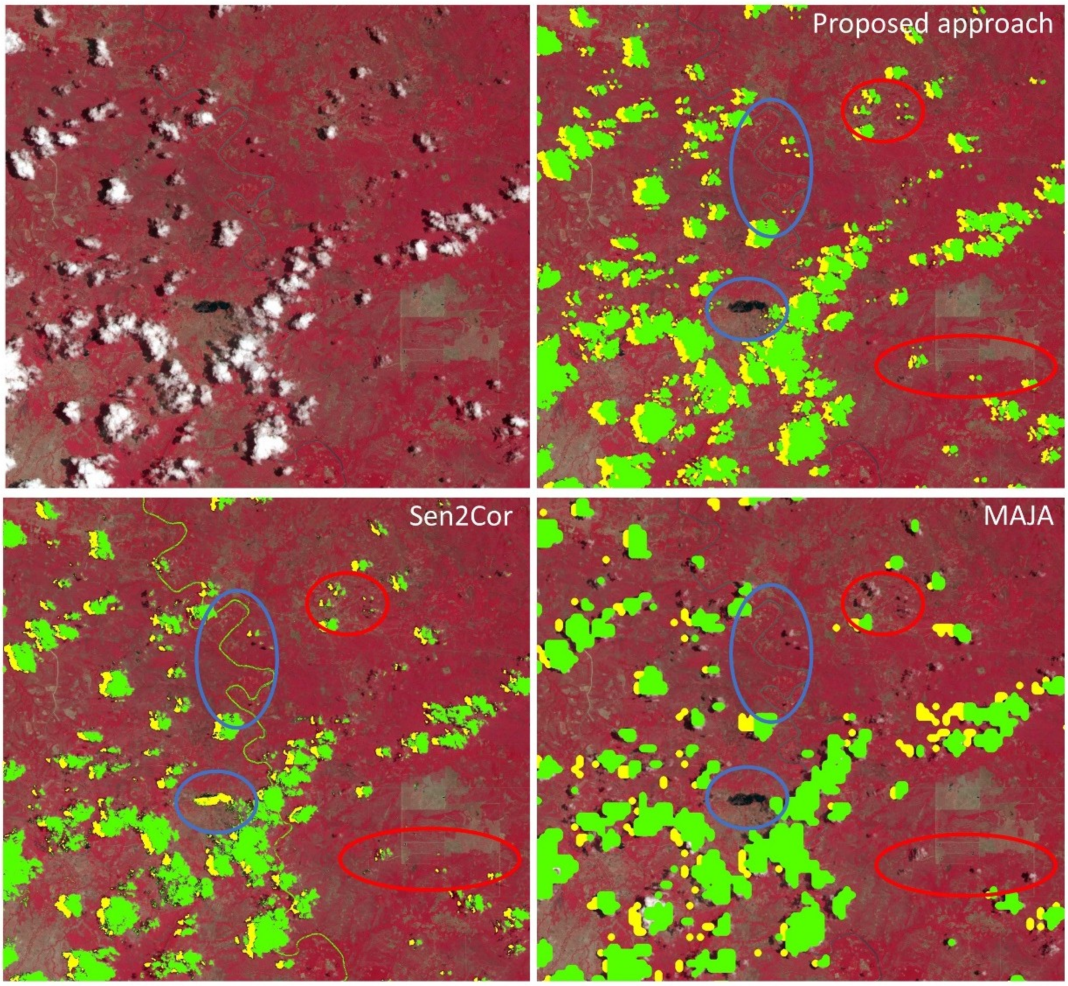


Figure 16. First test image of a Cambodia site. Zoom-in of the final 3-level cloud/cloud–shadow/others maps generated by the three algorithms under comparison, where class "others" is overlaid with the input Sentinel-2 A Multi-Spectral Instrument (MSI) Level 1C image, radiometrically calibrated into TOARF values and depicted in false-colors: monitor-typical RGB channels are selected as R = Near InfraRed (NIR) channel, G = Visible Red channel, B = Visible Green channel. Histogram stretching is applied for visualization purposes. Output class cloud is shown in a green pseudocolor, class cloud–shadow in a yellow pseudocolor. Based on qualitative photointerpetation, Sen2Cor appears to underestimate cloud–shadows, although some water areas are misclassified as cloud–shadows. In addition, some river/river beds are misclassified as clouds. These two cases of cloud false positives and cloud–shadow false positives are highlighted in blue circles. MAJA overlooks some clouds small in size (in relative terms), as highlighted in red circles.

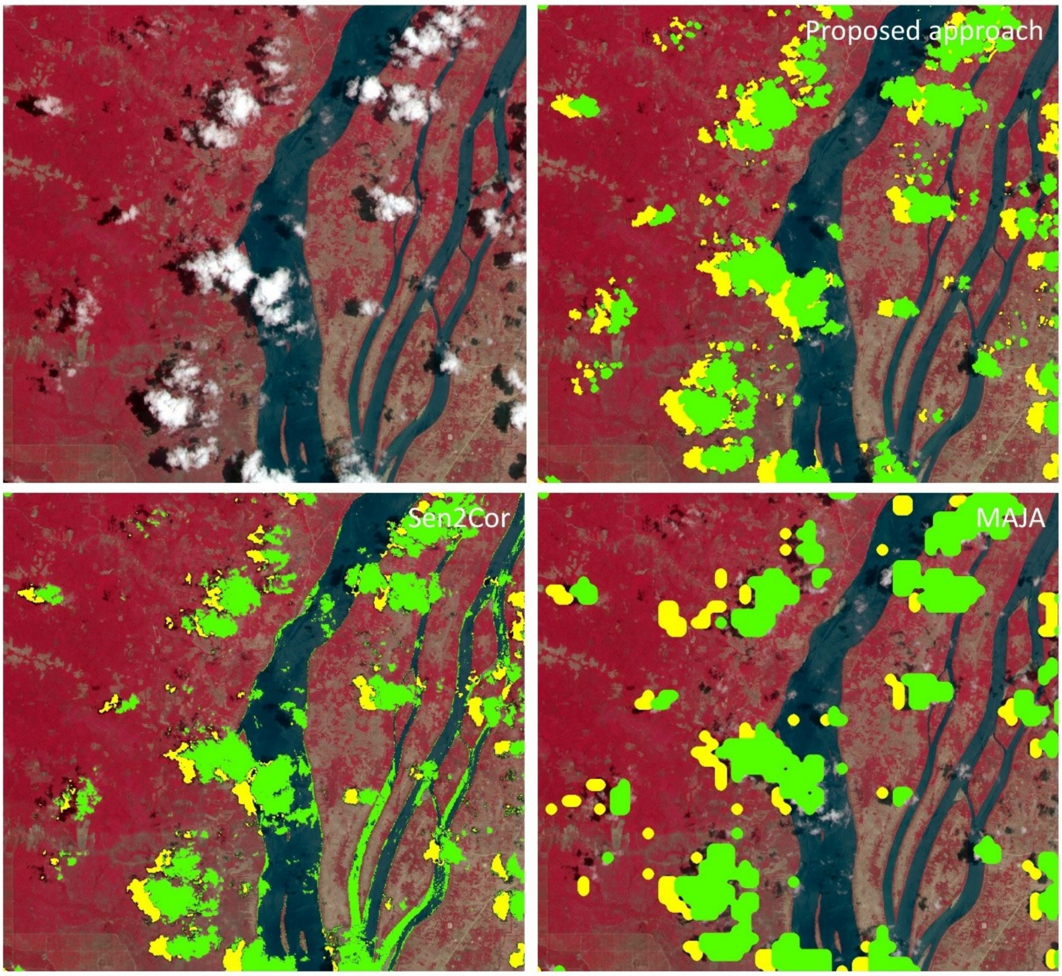


Figure 17. First test image of a Cambodia site. Extra zoom-in of the final 3-level cloud/cloud–shadow/others maps generated by the three algorithms under comparison, where class "others" is overlaid with the input Sentinel-2 A Multi-Spectral Instrument (MSI) Level 1C image, radiometrically calibrated into TOARF values and depicted in false-colors: monitor-typical RGB channels are selected as R = Near InfraRed (NIR) channel, G = Visible Red channel, B = Visible Green channel. Histogram stretching is applied for visualization purposes. Output class cloud is shown in a green pseudocolor, class cloud–shadow in a yellow pseudocolor. Based on qualitative photointerpetation, Sen2Cor appears to underestimate cloud–shadows, although some detected cloud–shadows are false positives because of misclassified water areas. In addition, some river/river beds are misclassified as clouds. To reduce false positives in cloud–shadow detection, MAJA adopts a multi-date approach. Nevertheless, MAJA misses some instances of cloud-over-water. Overall, MAJA cloud/cloud–shadow results look more "blocky" (affected by artifacts in localizing true boundaries of target image-objects).

Table 4. First test image of a Cambodia site. Comparison of cloud/cloud–shadow total areas detected by the three tested algorithms.

	Cloud area (ha)	Cloud–shadow (ha)
AutoCloud+	103811	49368
Sen2Cor*	82589	20296
MAJA	109373	25896

*including all cloud probability classes

Final cloud/cloud–shadow output maps generated from the second Sentinel-2 test image by the single-date AutoCloud+ and single-date Sen2Cor computer programs are shown in Figure 18, zoomed-in in Figure 19. Table 5 reports on the total area of clouds and cloud–shadows detected by the two algorithms of interest in the second test image.

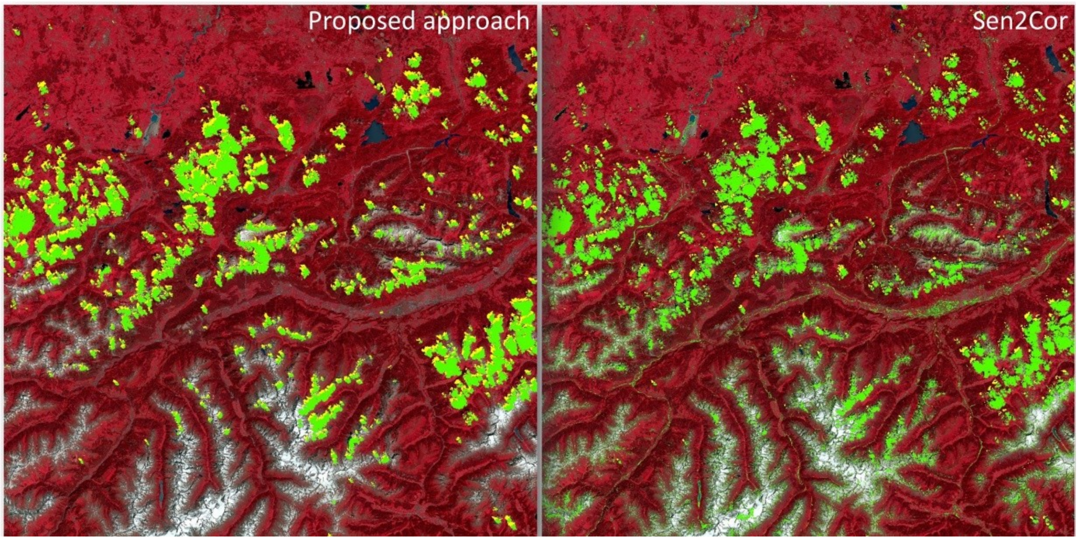


Figure 18. Second test image of an Alpine site. Final 3-level cloud/cloud–shadow/others maps generated by the two algorithms under comparison, specifically, single-date AutoCloud+ and single-date Sen2Cor, where class "others" is overlaid with the input Sentinel-2 B Multi-Spectral Instrument (MSI) Level 1C image, radiometrically calibrated into TOARF values and depicted in false-colors: monitor-typical RGB channels are selected as R = Near InfraRed (NIR) channel, G = Visible Red channel, B = Visible Green channel. Histogram stretching is applied for visualization purposes. Output class cloud is shown in a green pseudocolor, class cloud–shadow in a yellow pseudocolor.

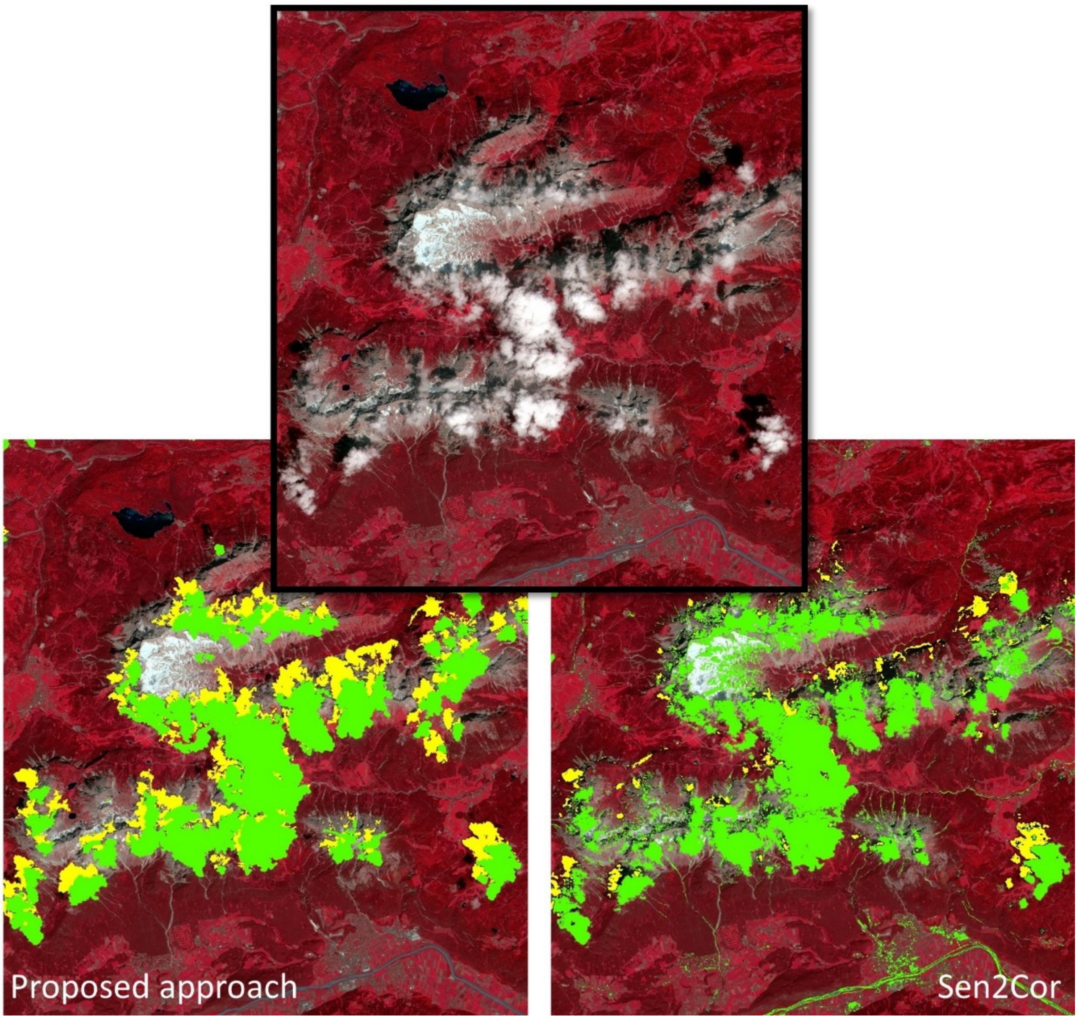


Figure 19. Second test image of an Alpine site. Zoom-in of the final 3-level cloud/cloud–shadow/others maps generated by the two algorithms under comparison, where class "others" is overlaid with the input Sentinel-2 A Multi-Spectral Instrument (MSI) Level 1C image, radiometrically calibrated into TOARF values and depicted in false-colors: monitor-typical RGB channels are selected as R = Near InfraRed (NIR) channel, G = Visible Red channel, B = Visible Green channel. Histogram stretching is applied for visualization purposes. Output class cloud is shown in a green pseudocolor, class cloud–shadow in a yellow pseudocolor. Based on qualitative photointerpetation, Sen2Cor appears to underestimate cloud–shadows, while some detected clouds are false positives because of misclassified water in river/river beds or misclassified snow/ice areas. AutoCloud+ avoids most of these cloud false positives and cloud–shadow false negatives by means of 2D spatial reasoning, specifically, by means of physical model-based modeling of cloud/cloud–shadow 2D shape properties and of bidirectional cloud/cloud–shadow 2D spatial relationships in the image-domain.

Table 5. Second test image of an Alpine site. Comparison of cloud/cloud–shadow total areas detected by the two tested algorithms.

	Cloud area (ha)	**Cloud–shadow (ha)**
AutoCloud+	106667	37905
Sen2Cor*	75180	14526

*including all cloud probability classes

6. Discussion

For the first Sentinel-2 test image of a Cambodia site, qualitative photointerpretation of the AutoCloud+ intermediate output products shown in Figure 14, in combination with quantitative vector quantization (VQ) error maps automatically generated as output by both the SIAM and RGBIAM software toolboxes (not shown), confirmed the validity of the SIAM and RGBIAM software solutions for color space quantization into categorical color names, in line with the existing literature [11,83,84,121], where SIAM was validated at continental scale on multi-year annual Landsat image mosaics. At first glance, a qualitative photointerpretation of the output 3-level cloud/cloud–shadow/others thematic maps generated by the three AutoCloud+, Sen2Cor and MAJA algorithms under comparison, shown in Figure 15, appears satisfactory. No cloud, large in size (in relative terms), is missed by any algorithm of interest. Actually, detected clouds appear somehow overestimated rather than underestimated in AutoCloud+ and MAJA. On a second glance, differences between the three approaches become notable. Table 4 reports the total area of extracted clouds and cloud–shadows. AutoCloud+ and MAJA score similar values for the cloudy total area (~103 000 ha / 109 000 ha), but the former scores twice as much in the total amount of cloud–shadow areas (~50 000 ha). The Sen2Cor results reveal the lowest total area values (~ 82 500 ha / 20 000 ha) for both classes cloud and cloud–shadow respectively. These summary statistics are confirmed by visual analysis of zoomed-in subsets of Figure 15, as shown in Figure 16 and Figure 17. AutoCloud+ and MAJA appear to dilate (grow, on purpose) clouds at cloud boundaries to include an aura of thin cloud regions, if any, which are typically ignored by Sen2Cor. In this spatial neighboring analysis, MAJA results look more "blocky", i.e., affected by artifacts in localizing perceptually true boundaries of target image-objects, whereas the AutoCloud+ approach "naturally" follows the shape of clouds as they are perceived by a human photo interpreter. Sen2Cor largely underestimates cloud–shadows, although some water areas are misclassified as cloud–shadow, see Figure 16. In addition, some river/river beds are misclassified as clouds, see Figure 16 and Figure 17. To avoid such false positives in cloud and cloud–shadow detection affecting Sen2Cor, MAJA adopts a multi-temporal approach. Nevertheless, MAJA misses some instances of clouds small in size (in relative terms), which are detected quite well by the other two algorithms, see Figure 16. In addition, MAJA misses some instances of cloud-over-water, see Figure 17. Overall, on closer inspection of the first test case, single-date AutoCloud+ scores qualitatively "high" in cloud and cloud–shadow mapping accuracy, both affected by few false positives and few false negatives. Hence, single-date AutoCloud+ is considered superior to single-date Sen2Cor in mapping accuracy and superior to multi-date MAJA in mapping accuracy as well as ease-of-use, timeliness (from EO data collection to EO data-derived VAPS generation) and costs in computer power, since the former requires less input data to run.

For the second Sentinel-2 test image of an Alpine site, qualitative photointerpretation of the output 3-level cloud/cloud–shadow/others thematic maps generated by the two AutoCloud+ and Sen2Cor algorithms under comparison, shown in Figure 18 and zoomed-in Figure 19, confirm conclusions drawn from the first test case. Sen2Cor appears to underestimate cloud–shadows, while some detected clouds are false positives because of misclassified water in river/river beds or misclassified snow/ice areas. AutoCloud+ avoids most of these cloud false positives and cloud–shadow false negatives. Image-wide statistics collected in Table 5 confirm that AutoCloud+ tends to detect 30% to 40% more cloud-affected areas (in ha) and 50% more areas affected by cloud–shadow phenomena than Sen2Cor.

To be validated in future works at large spatial scale and multiple time samples according to the GEO-CEOS *Val* guidelines [105], these preliminary results confirm the well-groundedness of our working hypothesis (see Figure 4 and Equation (1) in the Part 1 of this paper; otherwise, refer to the further Section 7 in the present Part 2, where such a working hypothesis is reported for the sake of completeness) in the multi-disciplinary domain of cognitive science (see Figure 2 in the Part 1), starting from the AutoCloud+ project requirements specification, summarized in Section 3 at the levels of system understanding known as information/knowledge representation, system architecture, algorithm and implementation (see Section 1).

AutoCloud+ process quality indicators belonging to the mDMI set of EO OP-Q^{2}Is proposed in Section 2 of the Part 1, but not investigated in this experimental session are robustness to changes in input data, at large spatial scale and time span in agreement with the GEO-CEOS *Val* guidelines [105], and scalability to changes in imaging sensor's spatial and spectral resolutions. In agreement with theory (refer to Section 3), AutoCloud+ is expected to score "high" in all EO OP-Q^{2}Is proposed in Section 2 of the Part 1.

7. Conclusions

The overarching goal of this research and technological development (RTD) study is to contribute toward filling an analytic and pragmatic information gap from multi-sensor, multi-temporal and multi-angular Earth observation (EO) *big image data cubes* into timely, comprehensive and operational EO data-derived value-adding information products and services (VAPS), in compliance with the intergovernmental Group on Earth Observations (GEO)'s visionary goal of a Global EO System of Systems (GEOSS) [2,3], unaccomplished to date. Main contributions of this paper pertain to the multi-disciplinary domain of cognitive science [27–32] (see Figure 2 in the Part 1), encompassing artificial general intelligence (AI) as *superset-of* computer vision (CV), i.e., dependence relationship '[AI ⊃ CV ⊃ Earth observation (EO) image understanding (EO-IU)] → cognitive science' holds in symbols of the standard Unified Modeling Language (UML) for graphical modeling of object-oriented software [33], where symbol '→' means *part-of* dependence, pointing from the supplier to the client, whereas symbol '⊃' means *subset-of* relationship (with inheritance), pointing from the superset to the subset.

For the sake of readability this paper is divided in two. To highlight the importance of a "universal" AutoCloud+ CV software system for cloud and cloud–shadow quality layers detection in multi-source, multi-angular and multi-temporal EO multi-spectral (MS) *big image data cubes*, the Part 1 presents AutoCloud+ in a broad context of systematic ESA EO Level 2 product generation at the ground segment [16,17], within a "seamless chain of innovation" needed for a new era of Space Economy 4.0 [4]. Provided with a relevant survey value, the Part 1 critically reviews the open problem of systematic ESA EO Level 2 product generation at the three higher (more abstract) levels of understanding of an information processing system proposed by Marr, specifically, outcome and process requirements specification, information/knowledge representation and system design (architecture) [11–15,134,135]. Typically considered the linchpin of success of any information processing system [11–15], these abstract levels of understanding make the critical review proposed in the Part 1 not alternative, but complementary to surveys on EO-IU system solutions typically presented in the remote sensing (RS) literature, such as [136–138], focused exclusively on the Marr two lowest levels of understanding, specifically, algorithm and implementation. Subsequent to the Part 1, the present Part 2 (proposed as Supplementary Materials) presents and discusses an original "universal" AutoCloud+ CV software system instantiation at the Marr five levels of understanding of an information processing system.

Original contributions and main conclusions of this two-part RTD study are summarized below.

Conceptual in nature and pertaining to the interdisciplinary domain of cognitive science, see Figure 2 in the Part 1, the first original contribution of this work coincides with our working hypothesis, refer to Equation (1) and Figure 4 in the Part 1. In symbols of the standard UML for graphical modeling of object-oriented software [33], our working hypothesis is formulated as follows:

'Human vision → CV ⊃ EO-IU in operating mode ⊃ NASA EO Level 2 product → ESA EO Level 2 product → [EO-SCBIR + SEIKD = AI4DIAS] → GEO-GEOSS'. (3)

Equation (3) is a dependence relationship, equivalent to a first principle (axiom, postulate). It postulates that necessary-but-not-sufficient pre-condition for multi-sensor EO *big data cube* analytics is systematic generation of ESA EO Level 2 product [16,17], encompassing cloud and cloud–shadow quality layer detection in an operating mode (for definition of operating mode of an EO *big data* processing system, refer to Section 2 in the Part 1). In more detail, systematic ESA EO Level 2 product generation is regarded as necessary-but-not-sufficient pre-condition for developing timely, comprehensive and operational EO data-derived VAPS, such as semantic content-based image retrieval (SCBIR) + semantics-enabled information/knowledge discovery (SEIKD) = AI for Data and Information Access Services at the ground segment (AI4DIAS), in multi-sensor, multi-temporal and multi-angular EO *big data cubes*, as *part-of* the GEO-CEOS visionary goal of a GEOSS, never accomplished to date by the RS community. The dependence relationship (3) implies that no solution to the dependent open problem of GEOSS, including its still-unsolved (open) sub-problems of SCBIR and SEIKD, can be found until the necessary-but-not-sufficient pre-condition of CV ⊃ EO-IU in operating mode, specifically, systematic ESA EO Level 2 product generation, is fulfilled in advance.

As a corollary of working hypothesis (3), a dependence relationship '*vision* (encompassing both biological vision and computer vision, CV) ⊃ CV ⊃ ESA EO Level 2 product' implies that generation of an ESA EO Level 2 product from a MS image is a CV ⊃ EO-IU task. Synonym for scene-from-image reconstruction and understanding [25], *vision* is a cognitive *(information-as-data-interpretation*) problem [27] very difficult to solve because: (i) non-deterministic polynomial (NP)-hard in computational complexity [34,35], (ii) and inherently ill-posed in the Hadamard sense [25,26,36]. *Vision* is inherently ill-posed because affected by: (I) a 4D-to-2D data dimensionality reduction, from the 4D geospatial-temporal scene-domain to the (2D, planar) image-domain, and (II) a semantic information gap from ever-varying sub-symbolic sensory data (sensations) in the physical world to stable symbolic percepts in the mental model of the physical world (modeled world, world ontology, real-world model) [11,12,25,27,37–40], see Figure 6 in the Part 1. Since *vision* is an inherently ill-posed and NP-hard cognitive *(information-as-data-interpretation*) problem [27] and the dependence relationship '*vision* ⊃ CV ⊃ ESA EO Level 2 product' holds, then CV, in general, and ESA EO Level 2 product generation, in particular, are inherently ill-posed and NP-hard cognitive problems too. As such, they are very difficult to solve and require *a priori* knowledge in addition to sensory data to become better conditioned for numerical solution [41,42]. In addition, according to the working hypothesis where dependence 'ESA EO Level 2 product → [EO-SCBIR + SEIKD = AI4DIAS] → GEOSS' holds, the computational complexity of GEOSS ≥ (not inferior to) computational complexity of [SCBIR + SEIKD] ≥ computational complexity of ESA EO Level 2 product generation, which is inherently ill-posed and NP-hard. It means that '[EO-SCBIR + SEIKD = AI4DIAS] → GEOSS' too are cognitive problems inherently ill-posed and NP-hard [27], whose solution is not less difficult to reach than that of CV in operating mode, unaccomplished to date. This corollary highlights, first, the inherent complexity and relevance of ESA EO Level 2 product as information primitive (unit of information) necessary-but-not-sufficient for EO data-derived VAPS generation in a "seamless innovation chain" needed for a new era of Space 4.0 [4], see Figure 8 in the Part 1. Second, it justifies the present RTD work, whose specific goal is cloud and cloud–shadow quality layers detection in operating mode as necessary-but-not-sufficient pre-condition for systematic ESA EO Level 2 product generation at the ground segment [16,17] or space segment [139–141].

Our second original contribution is both conceptual and pragmatic in the definition of RS best practices (see Section 2 in the Part 1), which is the focus of efforts made by intergovernmental organizations such as GEO and the Committee on Earth Observation Satellites (CEOS). The ESA EO Level 2 information product definition is regarded as baseline information primitive (unit of information) suitable for an "augmented" EO Analysis Ready

Data (ARD) format specification, more restrictive (in terms of output product requirements specification) and more informative (in terms of physical and conceptual/semantic quality of numeric and categorical output products, respectively), but more difficult to be inferred from EO sensory data than existing U.S. Landsat ARD [5–9] and CEOS ARD for Land (CARD4L) [10] format definitions.

Our final contribution in filling the gap from EO *big data* to EO data-derived VAPS stems from the present Part 2, focused on the RTD and the preliminary quality assessment of an innovative AutoCloud+ CV software system, eligible for "universal" cloud and cloud–shadow quality layer detection in multi-sensor multi-angular EO single-date MS imagery, either radiometrically uncalibrated, such as those typically acquired by lightweight imaging sensors mounted on small satellites [78] or small unmanned aerial vehicles (UAVs) [79], provided with no on-board radiometric calibration subsystem, or radiometrically calibrated into top-of-atmosphere reflectance (TOARF), surface reflectance (SURF) or surface albedo values [130,142–144], in agreement with the GEO-CEOS Quality Accuracy Framework for EO (QA4EO) Calibration/Validation (*Cal/Val*) requirements [2,105].

It is noteworthy that joint (combined, inter-dependent) detection of cloud and cloud–shadow quality layers is a typical example of physical model-based cause–effect relationship, expected to be very difficult to solve by inductive machine learning-from-data algorithms, such as increasingly popular deep convolutional neural networks (DCNNs) [65], with special regard to DCNNs designed and trained end-to-end for semantic segmentation [70] and instance segmentation [71], rather than object detection [69], where image-objects are localized with bounding boxes and categorized into one-of-many categories. In general, inductive machine learning-from-data algorithms, including DCNNs [65,68–71], are suitable for learning complex correlations between input and output features [41,42], but capable of no inherent representation of causality [96,97], in agreement with the well-known dictum that correlation does not imply causation and vice versa [11,12,42,96–98].

In comparison with standard cloud/cloud–shadow software toolboxes available either open source or free of cost, such as the single-date multi-sensor Function of Mask (FMask) open source algorithm [66,67], the single-date single-sensor ESA Sentinel 2 (atmospheric, topographic and adjacency) Correction Prototype Processor (Sen2Cor), to be run free-of-cost on the user side [16,17,72], and the multi-date Multisensor Atmospheric Correction and Cloud Screening (MACCS)-Atmospheric/Topographic Correction (ATCOR) Joint Algorithm (MAJA), developed and run by CNES/ CESBIO/ DLR [18–20] (refer to Section 3 in the Part 1), AutoCloud+ features several degrees of novelty at the Marr five levels of understanding of an information processing system [11–15]. For example, at the levels of understanding of information/knowledge representation, system architecture and algorithm, in comparison with standard open source or free-of-cost approaches in operating mode (see Table 5 in the Part 1), AutoCloud+ is the sole hybrid (combined physical and statistical model-based) 2D image analysis (spatial context-sensitive and spatial topology-preserving, see Figure 16 in the Part 1) approach, provided with feedback loops in agreement with biological cognitive systems [11,29,35,47–52,55–57,145].

In a proof-of-concept conducted by qualitative photointerpretation of two Sentinel-2 test images, an AutoCloud+ prototypical software implementation outperformed the standard single-date sensor-specific ESA Sen2Cor and the multi-date multi-sensor CNES/ CESBIO/ DLR MAJA computer programs in terms of cloud and cloud–shadow mapping quality indicators, such as true positive, false positive, true negative and false negative occurrences.

With regard to a minimally dependent and maximally informative (mDMI) set of EO outcome and process (OP) quantitative quality indicators (Q^2Is), to be community-agreed upon for use by members of the RS community, proposed in Section 2 of the Part 1, potential advantages of the AutoCloud+ hybrid feedback 2D image analysis approach for cloud and cloud–shadow quality layers detection in operating mode are summarized below.

(1) The AutoCloud+ computer program is "fully automated".

i. It requires no system's free-parameter to be user-defined based on heuristics, unlike mainstream CV and EO-IU algorithms, such as inductive DCNNs, where *a priori* knowledge is encoded by design based on empirical criteria, a.k.a. trial-and error. Hence, the AutoCloud+ ease of use cannot be surpassed by alternative algorithms.

ii. It requires no supervised data set for inductive learning-from-data. Hence, its cost in manpower for collecting training data samples, typically expensive and difficult to gather, is zero. In addition, its timeliness from EO data collection to VAPS generation is reduced to computation time, because training time is zero.

(2) It employs a known hybrid feedback CV system design (architecture, see Figure 2) whose low-level CV software units (modules) have been implemented to a large degree in operating mode and tested/validated in previous works. For example, refer to the Satellite Image Automatic Mapper™ (SIAM™) [13–15,80,83,84,111–116] and the RGB Image Automatic Mapper™ (RGBIAM™) [13,121,122], two lightweight computer programs for color (hyper)space (hyper)polyhedralization (partitioning) into a static vocabulary of color names, superpixel detection and vector quantization (VQ) quality assessment. Hence, the RTD costs in manpower of the AutoCloud+ software system have been reduced by a great deal, benefitting from the engineering principles of modularity, hierarchy and regularity considered mandatory by structured system design to guarantee scalability [104].

(3) In line with biological vision [11,13,35,47–63], the AutoCloud+ CV subsystem exploits dominant spatial topological and spatial non-topological information components in combination with secondary color information [25], discretized into color names, according to a convergence-of-evidence approach consistent with symbolic human reasoning, where potentially weak independent sources of evidence are typically combined to infer strong conjectures [11,12,25,39]. For example, in two Sentinel-2 test images selected for proof-of-concept, the single-date AutoCloud+ mapping accuracy, in both cloud and cloud–shadow detection and in terms of true positive, false positive, true negative and false negative estimates, scored better than that of standard software toolboxes, such as the single-date single-sensor ESA Sen2Cor and the multi-date multi-sensor CNES/ CESBIO/ DLR MAJA. In addition to showing a superior mapping accuracy, supported by theory (refer to Section 3 in the present Part 2), AutoCloud+ is expected to score "high" also in terms of robustness to changes in input data acquired across space, time and sensors, and in terms of scalability to changes in MS imaging sensor's spatial and spectral specifications, encompassing MS imagery either radiometrically uncalibrated or radiometrically calibrated into TOARF, SURF or surface albedo values.

To underpin the operational readiness of the AutoCloud+ software toolbox in support of systematic ESA EO Level 2 product generation at the ground segment or space segment [139–141], a complete AutoCloud+ software transcodification into the C/C++ programming language and a software integration phase are planned for efficiency reasons, together with a GEO-CEOS stage 3 and stage 4 *Val* campaign to be conducted by independent means at large spatial extent and time span with multiple MS imaging sensors, in compliance with the GEO-CEOS *Val* guidelines [105] and the GEO-CEOS QA4EO *Cal/Val* requirements [2].

Author Contributions: Conceptualization, A.B. and D.T.; Methodology, A.B. and D.T.; Software, A.B. and D.T.; Validation, D.T.; Formal Analysis, A.B. and D.T.; Investigation, A.B. and D.T.; Resources, A.B. and D.T.; Data Curation, D.T.; Writing-Original Draft Preparation, A.B.; Writing-Review & Editing, A.B. and D.T.; Visualization, A.B. and D.T.; Supervision, D.T.; Project Administration, D.T.; Funding Acquisition, D.T.

Funding: This research was funded in part by the Austrian Science Fund (FWF) through the Doctoral College GIScience (DK W1237- N23) and by the Austrian Research Promotion Agency (FFG) with regard to project Sentinel-2 Semantic Data Cube (Sen2cube).

Acknowledgments: Andrea Baraldi thanks Prof. Raphael Capurro for his hospitality, patience, politeness and open-mindedness. The authors wish to thank the Editor-in-Chief, Associate Editor and reviewers for their competence, patience and willingness to help.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

Acronyms

AI: Artificial general Intelligence
AI4DIAS: Artificial Intelligence for Data and Information Access Services (at the ground segment)
AI4Space: Artificial Intelligence for Space (segment)
ARD: Analysis Ready Data (format)
ATCOR: Atmospheric/Topographic Correction commercial 39oftware product
AVHRR: Advanced Very High Resolution Radiometer
BC: Basic Color
BIVRFTAB: Bivariate Frequency Table
Cal: Calibration
Cal/Val: Calibration and Validation
CBIR: Content-Based Image Retrieval
CEOS: Committee on Earth Observation Satellites
CESBIO: Centre d'Etudes Spatiales de la Biosphère
CFMask: C (programming language version of) Function of Mask
CLC: CORINE Land Cover (taxonomy)
CNES: Centre national d'études spatiales
CNN: Convolutional Neural Network
CORINE: Coordination of Information on the Environment
CV: Computer Vision
DCNN: Deep Convolutional Neural Network
DEM: Digital Elevation Model
DIAS: Data and Information Access Services
DLR: Deutsches Zentrum für Luft- und Raumfahrt (German Aerospace Center)
DN: Digital Number
DP: Dichotomous Phase (in the FAO LCCS taxonomy)
DRIP: Data-Rich, Information-Poor (syndrome)
EO: Earth Observation
EO-IU: EO Image Understanding
EO-IU4SQ: EO Image Understanding for Semantic Querying
ESA: European Space Agency
FAO: Food and Agriculture Organization
FIEOS: Future Intelligent EO imaging Satellites
Fmask: Function of Mask
GEO: Intergovernmental Group on Earth Observations
GEOSS: Global EO System of Systems
GIGO: Garbage In, Garbage Out principle of error propagation
GIS: Geographic Information System
GIScience: Geographic Information Science
GUI: Graphic User Interface
IGBP: International Global Biosphere Programme
IoU: Intersection over Union
IU: Image Understanding
LAI: Leaf Area Index

LC: Land Cover
LCC: Land Cover Change
LCCS: Land Cover Classification System (taxonomy)
LCLU: Land Cover Land Use
LEDAPS: Landsat Ecosystem Disturbance Adaptive Processing System
MAACS: Multisensor Atmospheric Correction and Cloud Screening
MAJA: Multisensor Atmospheric Correction and Cloud Screening (MACCS)-Atmospheric/Topographic Correction (ATCOR) Joint Algorithm
mDMI: Minimally Dependent and Maximally Informative (set of quality indicators)
MHP: Modular Hierarchical Phase (in the FAO LCCS taxonomy)
MIR: Medium InfraRed
MODIS: Moderate Resolution Imaging Spectroradiometer
MS: Multi-Spectral
MSI: (Sentinel-2) Multi-Spectral Instrument
NASA: National Aeronautics and Space Administration
NIR: Near InfraRed
NLCD: National Land Cover Data
NOAA: National Oceanic and Atmospheric Administration
NP: Non-Polynomial
OBIA: Object-Based Image Analysis
OGC: Open Geospatial Consortium
OP: Outcome (product) and Process
OP-Q^2I: Outcome and Process Quantitative Quality Index
QA4EO: Quality Accuracy Framework for Earth Observation
Q^2I: Quantitative Quality Indicator
RGB: monitor-typical Red-Green-Blue data cube
RMSE: Root Mean Square Error
RS: Remote Sensing
RTD: Research and Technological Development
SCBIR: Semantic Content-Based Image Retrieval
SCM: Scene Classification Map
SEIKD: Semantics-Enabled Information/Knowledge Discovery
Sen2Cor: Sentinel 2 (atmospheric, topographic and adjacency) Correction Prototype Processor
SIAM™: Satellite Image Automatic Mapper™
STRATCOR: Stratified Topographic Correction
SURF: Surface Reflectance
TIR: Thermal InfraRed
TM (superscript): (non-registered) Trademark
Tmask: Temporal Function of Mask
TOA: Top-Of-Atmosphere
TOARD: TOA Radiance
TOARF: TOA Reflectance
UAV: Unmanned Aerial Vehicle
UML: Unified Modeling Language
USGS: US Geological Survey
Val: Validation
VAPS: Value-Adding information Products and Services
VQ: Vector Quantization
WGCV: Working Group on Calibration and Validation

References

1. Yang, C.; Huang, Q.; Li, Z.; Liu K.; Hu, F. Big Data and cloud computing: Innovation opportunities and challenges. *Int. J. Digit. Earth* **2017**, *10*, 13–53, doi:10.1080/17538947.2016.1239771.
2. Group on Earth Observation/Committee on Earth Observation Satellites (GEO-CEOS). A Quality Assurance Framework for Earth Observation, Version 4.0. 2010. Available online: http://qa4eo.org/docs/QA4EO_Principles_v4.0.pdf (accessed on 17 Nov. 2018).
3. Group on Earth Observation (GEO). The Global Earth Observation System of Systems (GEOSS) 10-Year Implementation Plan. 2005. Available online: http://www.earthobservations.org/docs/10-Year%20Implementation%20Plan.pdf (accessed on 19 January 2012).
4. Mazzuccato M.; Robinson, D. Market Creation and the European Space Agency. European Space Agency (ESA) Report. 2017. Available online: https://marianamazzucato.com/wp-content/uploads/2016/11/Mazzucato_Robinson_Market_creation_and_ESA.pdf (accessed on 17 Nov. 2018).
5. U.S. Geological Survey (USGS). U.S. Landsat Analysis Ready Data (ARD). 2018. Available online: https://landsat.usgs.gov/ard (accessed on 15 July 2018).
6. U.S. Geological Survey (USGS). U.S. Landsat Analysis Ready Data (ARD) Data Format Control Book (DFCB) Version 4.0. January 2018. Available online: https://landsat.usgs.gov/sites/default/files/documents/LSDS-1873_US_Landsat_ARD_DFCB.pdf (accessed on 15 July 2018).
7. Dwyer, J.; Roy, D.; Sauer, B.; Jenkerson, C.; Zhang, H.; Lymburner, L. Analysis Ready Data: Enabling Analysis of the Landsat Archive. *Remote Sens.* **2018**, *10*, 1363, 1–19.
8. National Aeronautics and Space Administration (NASA). Harmonized Landsat/Sentinel-2 (HLS) Project. 2018. Available online: https://hls.gsfc.nasa.gov (accessed on 20 August 2018).
9. Helder, D.; Markham, B.; Morfitt, R.; Storey, J.; Barsi, J.; Gascon, F.; Clerc, S.; LaFrance, B.; Masek, J.; Roy, D.; et al. Observations and recommendations for the calibration of Landsat 8 OLI and Sentinel 2MSI for improved data interoperability. *Remote Sens.* **2018**, *10*, 1340, 1–29.
10. Committee on Earth Observation Satellites (CEOS). CEOS Analysis Ready Data–CEOS Analysis Ready Data for Land (CARD4L) Products. 2018. Available online: http://www.ceos.org/ard/ (accessed on 4 May 2018).
11. Baraldi, A. Pre-Processing, Classification and Semantic Querying of Large-Scale Earth Observation Spaceborne/Airborne/Terrestrial Image Databases: Process and Product Innovations. Ph.D. Thesis, Agricultural and Food Sciences, Department of Agricultural Sciences, University of Naples "Federico II", Naples, Italy, 2017. Available online: http://www.fedoa.unina.it/11724/1/TesiDottorato_Unina_XXIX_AB_EO-IU4SQ_v4_RGB_10April2017.pdf (accessed on 30 January 2018).
12. Sonka, M.; Hlavac, V.; Boyle, R. *Image Processing, Analysis and Machine Vision*; Chapman & Hall: London, UK, 1994.
13. Marr, D. *Vision*; Freeman and C: New York, NY, USA, 1982.
14. Baraldi, A.; Boschetti, L. Operational automatic remote sensing image understanding systems: Beyond Geographic Object-Based and Object-Oriented Image Analysis (GEOBIA/GEOOIA)–Part 1: Introduction. *Remote Sens.* **2012**, *4*, 2694–2735.
15. Baraldi, A.; Boschetti, L. Operational automatic remote sensing image understanding systems: Beyond Geographic Object-Based and Object-Oriented Image Analysis (GEOBIA/GEOOIA)–Part 2: Novel system architecture, information/knowledge representation, algorithm design and implementation. *Remote Sens.* **2012**, *4*, 2768–2817.
16. European Space Agency (ESA). *Sentinel-2 User Handbook, Standard Document*; ESA: Paris, France, 2015.
17. Deutsches Zentrum für Luft-und Raumfahrt e.V. (DLR); VEGA Technologies. *Sentinel-2 MSI–Level 2A Products Algorithm Theoretical Basis Document*; Document S2PAD-ATBD-0001; European Space Agency: Paris, France, 2011.
18. Hagolle, O.; Huc, M.; Desjardins, C.; Auer, S.; Richter, R. MAJA Algorithm Theoretical Basis Document. 2017. Available online: https://zenodo.org/record/1209633#.W2ffFNIzZaQ (accessed on 17 Nov. 2018).
19. Hagolle, O.; Rouquié, B.; Desjardins, C.; Makarau, A.; Main-knorn, M.; Rochais, G.; Pug, B. Recent Advances in Cloud Detection and Atmospheric Correction Applied to Time Series of High

Resolution Images. RAQRS. 19 September 2017. Available online: https://www.researchgate.net/profile/Olivier_Hagolle2/publication/320402521_Recent_advances_in_cloud_detection_and_atmospheric_correction_applied_to_time_series_of_high_resolution_images/links/59eeee074585154350e83669/Recent-advances-in-cloud-detection-and-atmospheric-correction-applied-to-time-series-of-high-resolution-images.pdf (accessed on 8 August 2018).
20. Hagolle, O.; Huc, M.; Pascual, D.V.; Dedieu, G.A multi-temporal method for cloud detection, applied to FORMOSAT-2, VENμS, LANDSAT and SENTINEL-2 images. *Remote Sens. Environ.* **2010**, *114*, 1747–1755, doi:10.1016/j.rse.2010.03.002.
21. Richter, R.; Schläpfer, D. Atmospheric/Topographic Correction for Satellite Imagery–ATCOR-2/3 User Guide. Version 8.2 BETA. 2012. Available online: http://www.dlr.de/eoc/Portaldata/60/Resources/dokumente/5_tech_mod/atcor3_manual_2012.pdf (accessed on 12 April 2013).
22. Richter, R.; Schläpfer, D. Atmospheric/Topographic Correction for Airborne Imagery–ATCOR-4 User Guide, Version 6.2 BETA. 2012. Available online: http://www.dlr.de/eoc/Portaldata/60/Resources/dokumente/5_tech_mod/atcor4_manual_2012.pdf (accessed on 12 April 2013).
23. Dorigo, W.; Richter, R.; Baret, F.; Bamler, R.; Wagner, W. Enhanced automated canopy characterization from hyperspectral data by a novel two step radiative transfer model inversion approach. *Remote Sens.* **2009**, *1*, 1139–1170.
24. Schläpfer, D.; Richter, R.; Hueni, A. Recent developments in operational atmospheric and radiometric correction of hyperspectral imagery. In Proceedings of the 6th EARSeL SIG IS Workshop, Tel Aviv, Israel, 16–19 March 2009. Available online: http://www.earsel6th.tau.ac.il/~earsel6/CD/PDF/earsel-PROCEEDINGS/3054%20Schl%20pfer.pdf (accessed on14 July 2012)
25. Matsuyama, T.; Hwang, V.S. *SIGMA–A Knowledge-Based Aerial Image Understanding System*; Plenum Press: New York, NY, 1990.
26. Bertero, M.; Poggio, T.; Torre, V. Ill-posed problems in early vision. *Proc. IEEE* **1988**, *76*, 869–889.
27. Capurro, R.; Hjørland, B. The concept of information. *Annu. Rev. Inf. Sci. Technol.* **2003**, *37*, 343–411.
28. Serra, R.; Zanarini, G. *Complex Systems and Cognitive Processes*; Springer-Verlag: Berlin, Germany, 1990.
29. Parisi, D. *La Scienza Cognitive tra Intelligenza Artificiale e vita Artificiale, in Neurosceinze e Scienze dell'Artificiale: Dal Neurone all'Intelligenza*; Patron Editore: Bologna, Italy, 1991.
30. Miller, G.A. The cognitive revolution: A historical perspective. *Trends Cogn. Sci.*, **2003**, *7*, 141–144.
31. Varela, F.J.; Thompson, E.; Rosch, E. *The Embodied Mind: Cognitive Science and Human Experience*; MIT Press: Cambridge, MA, USA, 1991.
32. Capra, F.; Luisi, P.L. *The Systems View of Life: A Unifying Vision*; Cambridge University Press: Cambridge, UK, 2014.
33. Fowler, M. *UML Distilled*, 3rd ed.; Addison-Wesley: Boston, MA, USA, 2003.
34. Frintrop, S. Computational visual attention. In *Computer Analysis of Human Behavior, Advances in Pattern Recognition*; Salah, A.A., Gevers, T., Eds.; Springer: Berlin, Germany, 2011.
35. Tsotsos, J.K. Analyzing vision at the complexity level. *Behav. Brain Sci.* **1990**, *13*, 423–469.
36. Hadamard, J. Sur les problemes aux derivees partielles et leur signification physique. *Princet. Univ. Bull.* **1902**, *13*, 49–52.
37. Fonseca, F.; Egenhofer, M.; Agouris, P.; Camara, G. Using ontologies for integrated geographic information systems, *Trans. GIS* **2002**, *6*, 231−257.
38. Growe, S. Knowledge-based interpretation of multisensor and multitemporal remote sensing images. *Int. Arch. Photogramm. Remote Sens*. **1999**, *32*, 71.
39. Laurini, R.; Thompson, D. *Fundamentals of Spatial Information Systems*; Academic Press: London, UK, 1992.
40. Sowa, J. *Knowledge Representation: Logical, Philosophical, and Computational Foundations*; Brooks Cole Publishing Co.: Pacific Grove, CA, USA, 2000.
41. Bishop, C.M. *Neural Networks for Pattern Recognition*; Clarendon: Oxford, UK, 1995.
42. Cherkassky, V.; Mulier, F. *Learning from Data: Concepts, Theory, and Methods*; Wiley: New York, NY, USA, 1998.

43. Tobler, W.R. A computer movie simulating urban growth in the Detroit Region. *Econ. Geogr.* **1970**, *46*, 234–240.
44. Longley, P.A.; Goodchild, M.F.; Maguire, D.J.; Rhind, D.W. *Geographic Information Systems and Science*, 2nd ed.; Wile: New York, NY, USA, 2005.
45. Blaschke, T.; Lang, S. Object based image analysis for automated information extraction-a synthesis. In Proceedings of the Measuring the Earth II ASPRS Fall Conference, San Antonio, CA, USA, 6–10 November 2006; pp. 6–10.
46. Blaschke, T.; Hay, G.J.; Kelly, M.; Lang, S.; Hofmann, P.; Addink, E.; Feitosa, R.Q.; van der Meer, F.; van der Werff, H.; van Coillie, F.; et al. Geographic object-based image analysis–towards a new paradigm. *ISPRS J. Photogram. Remote Sens.* **2014**, *87*, 180–191.
47. DiCarlo, J. The Science of Natural Intelligence: Reverse Engineering Primate Visual Perception. Keynote. CVPR17 Conference, 2017. Available online: https://www.youtube.com/watch?v=ilbbVkIhMgo (accessed on 5 January 2018).
48. du Buf, H.; Rodrigues, J. Image morphology: From perception to rendering. In *IMAGE–Computational Visualistics and Picture Morphology*; The University of Algarve: Faro, Portugal, 2007.
49. Kosslyn, S.M. *Image and Brain*; MIT Press: Cambridge, MA, USA, 1994.
50. Serre, T.; Wolf, L.; Bileschi, S.; Riesenhuber, M.; Poggio, T. Robust object recognition with cortex-like mechanisms. *IEEE Trans. Pattern Anal. Mach. Intell.* **2007**, *29*, 411–426.
51. Sylvester, J.; Reggia, J. Engineering neural systems for high-level problem solving. *Neural Netw.* **2016**, *79*, 37–52.
52. Vecera, S.; Farah, M. Is visual image segmentation a bottom-up or an interactive process? *Percept. Psychophys*. **1997**, *59*, 1280–1296.
53. Burt, P.; Adelson, E. The laplacian pyramid as a compact image code. *IEEE Trans. Commun.* **1983**, *31*, 532–540.
54. Jain, A.; Healey, G. A multiscale representation including opponent color features for texture recognition. *IEEE Trans. Image Process.* **1998**, *7*, 124–128.
55. Slotnick, S.D.; Thompson, W.L.; Kosslyn, S.M. Visual mental imagery induces retinotopically organized activation of early visual areas. *Cereb. Cortex* **2005**, *15*, 1570–1583.
56. Mély, D.; Linsley, D.; Serre, T. Complementary surrounds explain diverse contextual phenomena across visual modalities. *Psychol. Rev.* Sep 20, 2018.
57. Teaching Computers to See Optical Illusions. Available online: https://neurosciencenews.com/optical-illusions-neural-network-ai-9901/ (accessed on 1 October 2018).
58. Mason, C.; Kandel, E.R. Central Visual Pathways. In *Principles of Neural Science*; Kandel, E., Schwartz, J., Eds.; Appleton and Lange: Norwalk, CT, USA, 1991; pp. 420–439.
59. Gouras, P. Color Vision. In *Principles of Neural Science*; Kandel, E., Schwartz, J., Eds.; Appleton and Lange: Norwalk, CT, USA, 1991; pp. 467–479.
60. Kandel, E.R. Perception of Motion, Depth and Form. In *Principles of Neural Science*; Kandel, E., Schwartz, J., Eds.; Appleton and Lange: Norwalk, CT, USA; 1991; pp. 441–466.
61. Wilson, H.R.; Bergen, J.R. A four mechanism model for threshold spatial vision. *Vis. Res*. **1979**, *19*, 19–32.
62. Hubel, D.; Wiesel, T. Receptive fields of single neurons in the cat's striate cortex. *J. Physiol.* **1959**, *148*, 574–591.
63. Wiesel, T.N.; Hubel, D.H. Spatial and chromatic interactions in the lateral geniculate body of the rhesus monkey. *J. Neurophys.* **1966**, *29*, 1115–1156.
64. Martinetz, T.; Berkovich, G.; Schulten, K. Topology representing networks. *Neural Netw.* **1994**, *7*, 507–522.
65. Cimpoi M.; Maji, S.; Kokkinos I.; Vedaldi, A. Deep filter banks for texture recognition, description, and segmentation. *Int. J. Comput. Vis.* **2014**, doi:10.1007/s11263-015-0872-3.
66. Zhu, Z.; Woodcock, C.E. Object-based cloud and cloud-shadow detection in Landsat imagery. *Remote Sens. Environ.* **2012**, *118*, 83–94.
67. Zhu, Z.; Wang, S.; Woodcock, C.E. Improvement and expansion of the Fmask algorithm: Cloud, cloud-shadow, and snow detection for Landsats 4–7, 8, and Sentinel 2 images. *Remote Sens. Environ.* **2015**, *159*, 269–277. doi:10.1016/j.rse.2014.12.014.

68. Bartoš, M. Cloud and Shadow Detection in Satellite Imagery. Master's Thesis, Computer Vision and Image Processing, Faculty of Electrical Engineering, Department of Cybernetics, Czech Technical University, Prague, Czech Republic, 2017.
69. Li, K.; Cheng, G.; Bu, S.; You, X. Rotation-Insensitive and Context-Augmented Object Detection in Remote Sensing Images. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56*, 2337–2348.
70. Long, J.; Shelhamer, E.; Darrell, T. Fully convolutional networks for semantic segmentation. *IEEE Trans Pattern Anal. Mach. Intell.* **2017**, *39*, 640–651.
71. He, K.; Gkioxari, G.; Dol'ar, P.; Girshick, R. Mask R-CNN. *arXiv* **2018**, arXiv:1703.06870v3.
72. Gascon, F.; Bouzinac, C.; Thépaut, O.; Jung, M.; Francesconi, B.; Louis, J.; Lonjou, V.; Lafrance, B.; Massera, S.; Gaudel-Vacaresse, A.; et al. Copernicus Sentinel-2A Calibration and Products Validation Status. *Remote Sens.* **2017**, *9*, 584. doi:10.3390/rs9060584.
73. Main-Knorn, M.; Louis, J.; Hagolle, O.; Müller-Wilm, U.; Alonso, K. The Sen2Cor and MAJA cloud masks and classification products. In Proceedings of the 2nd Sentinel-2 Validation Team Meeting, ESA-ESRIN, Frascati, Rome, Italy, 29–31 January 2018.
74. Foga, S.; Scaramuzza, P.; Guo, S.; Zhu, Z.; Dilley, R., Jr.; Beckmann, T.; Schmidt, G.; Dwyer, J.; Hughes, M.; Laue, B. Cloud detection algorithm comparison and validation for operational Landsat data products. *Remote Sens. Environ.* **2017**, *194*, 379–390.
75. U.S. Geological Survey (USGS). U.S. Landsat Analysis Ready Data (ARD) Artifacts. 2018. Available online: https://landsat.usgs.gov/us-landsat-ard-artifacts (accessed on 15 July 2018).
76. Baraldi, A. Automatic Spatial Context-Sensitive Cloud/Cloud-Shadow Detection in Multi-Source Multi-Spectral Earth Observation Images–AutoCloud+, Invitation to tender ESA/AO/1-8373/15/I-NB–VAE: Next Generation EObased Information Services. *arXiv* **2015**, arXiv:1701.04256. Available online: https://arxiv.org/ftp/arxiv/papers/1701/1701.04256.pdf (accessed on 8 January 2017).
77. Baraldi, A.; Tiede, D. AutoCloud+, a "universal" single-date multi-sensor physical and statistical model-based spatial context-sensitive cloud/cloud-shadow detector in multi-spectral Earth observation imagery. In Proceedings of the GEOBIA 2018, Montpellier, France, 18–22 June 2018.
78. Satellite Applications Catapult, Small Is the New Big–Nano/Micro-Satellite Missions for Earth Observation and Remote Sensing. White Paper. 2018. Available online: https://sa.catapult.org.uk/wp-content/uploads/2016/03/Small-is-the-new-Big.pdf (accessed on 17 Nov. 2018).
79. Small Drones Market by Type (Fixed-Wing, Rotary-Wing, Hybrid/Transitional), Application, MTOW (<5 kg, 5–25 kg, 25–150 kg), Payload (Camera, CBRN Sensors, Electronic Intelligence Payload, Radar), Power Source, and Region–Global Forecast to 2025. Available online: https://www.researchandmarkets.com/research/lkh233/small_drones?w=12 (accessed on 17 Nov. 2018).
80. Baraldi, A.; Gironda, M.; Simonetti, D. Operational two-stage stratified topographic correction of spaceborne multi-spectral imagery employing an automatic spectral rule-based decision-tree preliminary classifier. *IEEE Trans. Geosci. Remote Sens.* **2010**, *48*, 112–146.
81. Smeulders, A.; Worring, M.; Santini, S.; Gupta, A.; Jain, R. Content-based image retrieval at the end of the early years. *IEEE Trans. Pattern Anal. Mach. Intell.* **2000**, *22*, 1349–1380.
82. Tyagi, V. *Content-Based Image Retrieval Ideas, Influences, and Current Trends*; Springer: Singapore, 2017.
83. Baraldi, A.; Humber, M.L.; Tiede, D.; Lang, S. 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. *Cogent Geosci*. **2018**, 1467357, doi:10.1080/23312041.2018.1467357.
84. Baraldi, A.; Humber, M.L.; Tiede, D.; Lang, S. GEO-CEOS stage 4 validation of the Satellite Image Automatic Mapper lightweight computer program for ESA Earth observation Level 2 product generation–Part 2: Validation. *Cogent Geosci.* **2018**, 1467254, doi:10.1080/23312041.2018.1467254.
85. Nagao, M.; Matsuyama, T. *A Structural Analysis of Complex Aerial Photographs*; Plenum Press: New York, NY, USA, 1980.
86. Shackelford, K.; Davis, C.H. A combined fuzzy pixel-based and object-based approach for classification of high-resolution multispectral data over urban areas. *IEEE Trans. Geosci. Remote Sens.* **2003**, *41*, 2354–2363.

87. Shackelford, K.; Davis, C.H. A hierarchical fuzzy classification approach for high-resolution multispectral data over urban areas. *IEEE Trans. Geosci. Remote Sens.* **2003**, *41*, 1920–1932.
88. Castelletti, D.; Pasolli, L.; Bruzzone, L.; Notarnicola, C.; Demir, B. A novel hybrid method for the correction of the theoretical model inversion in bio/geophysical parameter estimation. *IEEE Trans. Geosci. Remote Sens*. **2016**, *54*, 4764–4774.
89. Berlin, B.; Kay, P. *Basic Color Terms: Their Universality and Evolution*; University of California: Berkeley, CA, USA, 1969.
90. Griffin, L.D. Optimality of the basic color categories for classification. *J. R. Soc. Interface* **2006**, *3*, 71–85.
91. Li, W.; Goodchild, M.F.; Church, R.L. An efficient measure of compactness for 2D shapes and its application in regionalization problems. *Int. J. Geogr. Inf. Sci.* 2013, 1–24.
92. Baraldi, A.; Soares, J.V.B. 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. Available online: https://arxiv.org/ftp/arxiv/papers/1701/1701.01941.pdf (accessed on 8 January 2017).
93. Julesz, B. Texton gradients: The texton theory revisited. In *Biomedical and Life Sciences Collection*; Springer, Berlin/Heidelberg, Germany, 1986; Volume 54.
94. Julesz, B.; Gilbert, E.N.; Shepp, L.A.; Frisch, H.L. Inability of humans to discriminate between visual textures that agree in second-order statistics–revisited. *Perception* **1973**, *2*, 391–405.
95. Yellott, J. Implications of triple correlation uniqueness for texture statistics and the Julesz conjecture. *Opt. Soc. Am.* **1993**, *10*, 777–793.
96. Marcus, G. Deep Learning: A Critical Appraisal. *arXiv* **2018**, arXiv:1801.00631. Available online: https://arxiv.org/ftp/arxiv/papers/1801/1801.00631.pdf (accessed on 16 January 2018).
97. Bupe, C. Is Deep Learning Fundamentally Flawed and Hitting a Wall? Was Gary Marcus Correct in Pointing out Deep Learning's Flaws? Quora 2018. Available online: https://www.quora.com/Is-Deep-Learning-fundamentally-flawed-and-hitting-a-wall-Was-Gary-Marcus-correct-in-pointing-out-Deep-Learnings-flaws (accessed on 7 August 2018).
98. Pearl, J. *Causality: Models, Reasoning and Inference*; Cambridge University Press: New York, NY, USA, 2009.
99. Zadeh, L.A. Fuzzy sets. *Inf. Control* **1965**, *8*, 338–353.
100. Open Geospatial Consortium (OGC) Inc. 2015. *OpenGIS® Implementation Standard for Geographic information–Simple Feature Access–Part 1: Common Architecture*; OGC: Wayland, MA, USA, 2015. Available online: http://www.opengeospatial.org/standards/is (accessed on 8 March 2015)
101. Shotton, J.; Winn, J.; Rother, C.; Criminisi, A. Texton-Boost for Image Understanding: Multi-Class Object Recognition and Segmentation by Jointly Modeling Texture, Layout, and Context. *Int. J. Comp. Vis.* **2009**, *81*, 2–23.
102. Wikipedia. Latent Variable. 2018. Available online: https://en.wikipedia.org/wiki/Latent_variable (accessed on 30 January 2018).
103. Hunt, N.; Tyrrell, S. Stratified Sampling. Coventry University. 2012. Available Online: http://www.coventry.ac.uk/ec/~nhunt/meths/strati.html (accessed on 3 Jan. 2012).
104. Lipson, H. Principles of modularity, regularity, and hierarchy for scalable systems. *J. Biol. Phys. Chem.* **2007**, *7*, 125–128.
105. Group on Earth Observation/Committee on Earth Observation Satellites (GEO-CEOS)—Working Group on Calibration and Validation (WGCV). 2015. Land Product Validation (LPV). Available online: http://lpvs.gsfc.nasa.gov/ (accessed on 20 March 2015).
106. Baraldi, A.; Tiede, D.; Sudmanns, M.; Belgiu, M.; Lang, S. Automated near real-time Earth observation Level 2 product generation for semantic querying. In Proceedings of the GEOBIA 2016, University of Twente Faculty of Geo-Information and Earth Observation (ITC), Enschede, The Netherlands, 14–16 September 2016.
107. Baraldi, A.; Tiede, D.; Sudmanns, M.; Lang, S. Systematic ESA EO Level 2 product generation as pre-condition to semantic content-based image retrieval and information/knowledge discovery in EO image databases. In Proceedings of the BiDS'17 2017 Conference on Big Data from Space, Toulouse, France, 28–30 March 2017.

108. Tiede, D.; Baraldi, A.; Sudmanns, M.; Belgiu, M.; Lang, S. Architecture and prototypical implementation of a semantic querying system for big earth observation image bases. *Eur. J. Remote Sens.* **2017**, *50*, 452–463, doi:10.1080/22797254.2017.1357432.
109. Augustin, H.; Sudmanns, M.; Tiede, D.; Baraldi, A. A semantic Earth observation data cube for monitoring environmental changes during the Syrian conflict. In Proceedings of the AGIT 2018, Salzburg, Austria, 3–6 July 2018; pp. 214–227, doi:10.1553/giscience2018_01_s214.
110. Sudmanns, M.; Tiede, D.; Lang, S.; Baraldi, A. Semantic and syntactic interoperability in online processing of big Earth observation data. *Int. J. Digit. Earth* **2018**, *11*, 95–112, doi:10.1080/17538947.2017.1332112.
111. Baraldi, A.; Boschetti, L.; Humber, M. Probability sampling protocol for thematic and spatial quality assessments of classification maps generated from spaceborne/airborne very high resolution images. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 701–760.
112. Baraldi, A.; Durieux, L.; Simonetti, D.; Conchedda, G.; Holecz, F.; Blonda, P. Automatic spectral rule-based preliminary classification of radiometrically calibrated SPOT-4/-5/IRS, AVHRR/MSG, AATSR, IKONOS/QuickBird/OrbView/GeoEye and DMC/SPOT-1/-2 imagery–Part I: System design and implementation. *IEEE Trans. Geosci. Remote Sens.* **2010**, *8*, 1299–1325.
113. Baraldi, A.; Durieux, L.; Simonetti, D.; Conchedda, G.; Holecz, F.; Blonda, P. Automatic spectral rule-based preliminary classification of radiometrically calibrated SPOT-4/-5/IRS, AVHRR/MSG, AATSR, IKONOS/QuickBird/OrbView/GeoEye and DMC/SPOT-1/-2 imagery–Part II: Classification accuracy assessment. *IEEE Trans. Geosci. Remote Sens.* **2010**, *48*, 1326–1354.
114. Baraldi, A.; Humber, M. 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 J. Sel. Top. Appl. Earth Obs. Remote Sens*. **2015**, *8*, 1307–1329.
115. Baraldi, A.; Humber, M.; Boschetti, L. 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 Sens*. **2013**, *5*, 5209–5264.
116. Baraldi, A.; Puzzolo, V.; Blonda, P.; Bruzzone, L.; Tarantino, C. Automatic spectral rule-based preliminary mapping of calibrated Landsat TM and ETM+ images. *IEEE Trans. Geosci. Remote Sens.* **2006**, *44*, 2563–2586.
117. Di Gregorio, A.; Jansen, L. *Land Cover Classification System (LCCS): Classification Concepts and User Manual*; FAO Corporate Document Repository; FAO: Rome, Italy, 2000. http://www.fao.org/DOCREP/003/X0596E/X0596e00.htm (accessed on 10 February 2015).
118. Gijsenij, A.; Gevers, T.; van de Weijer, J. Computational color constancy: Survey and experiments. *IEEE Trans. Image Proc.* **2010**, *20*, 2475–2489.
119. Finlayson, G.D.; Hordley, S.D.; Hubel, P.M. Color by correlation: A simple, unifying framework for color constancy. *IEEE Trans. Pattern Anal. Mach. Intell.* **2001**, *23*, 1209–1221.
120. Gevers, T.; Gijsenij, A.; van de Weijer, J.; Geusebroek, J-M. *Color in Computer Vision*; Wiley: Hoboken, NJ, USA, 2012.
121. Vo, A.-V.; Truong-Hong, L.; Laefer, D.F.; Tiede, D.; d'Oleire-Oltmanns, S.; Baraldi, A.; Shimoni, M. Processing of extremely high resolution LiDAR and optical data: Outcome of the 2015 IEEE GRSS Data Fusion Contest. Part–B: 3D contest. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens*. **2016**, *9*, 5560–5575.
122. Baraldi, A.; Tiede, D.; Lang, S. Automated Linear-Time Detection and Quality Assessment of Superpixels in Uncalibrated True- or False-Color RGB Images. *arXiv* **2017,** arXiv:1701.01940. Available online: https://arxiv.org/ftp/arxiv/papers/1701/1701.01940.pdf (accessed on 8 January 2017).
123. Dillencourt, M.B.; Samet, H.; Tamminen, M. A general approach to connected component labeling for arbitrary image representations, *J. Assoc. Comput. Mach.* **1992**, *39*, 253–280.
124. Pessoa, L. Mach Bands: How Many Models are Possible? Recent Experimental Findings and Modeling Attempts. *Vision Res*. **1996**, *36*, 3205–3227.

125. Baatz, M.; Schäpe, A. Multiresolution Segmentation: An Optimization Approach for High Quality Multi-Scale Image Segmentation. In *Angewandte Geographische Informationsverarbeitung XII*; Strobl, J., Ed.; Herbert Wichmann Verlag: Berlin, Germany, 2000; Volume 58, pp. 12–23.
126. Espindola, G.M.; Camara, G.; Reis, I.A.; Bins, L.S.; Monteiro, A.M. Parameter selection for region-growing image segmentation algorithms using spatial autocorrelation. *Int. J. Remote Sens.;* **2006**, *27*, 3035–3040.
127. Canny, J. A computational approach to edge detection. *IEEE Trans. Pattern Anal. Mach. Intell*. **1986**, *8*, 679–698.
128. Liang, S. *Quantitative Remote Sensing of Land Surfaces*; John Wiley and Sons: Hoboken, NJ, USA, 2004.
129. Purdue University. Perceptual Grouping. 2018. Available online: http://cs.iupui.edu/~tuceryan/research/ComputerVision/perceptual-grouping.html (accessed on 10 August 2018).
130. Schaepman-Strub, G.; Schaepman, M.E.; Painter, T.H.; Dangel, S.; Martonchik, J.V. Reflectance quantities in optical remote sensing—Definitions and case studies. *Remote Sens. Environ.* **2006**, *103*, 27–42.
131. Trimble. *eCognition® Developer 9.0 Reference Book*; Trimble: Sunnyvale, CA, USA, 2015.
132. European Commission (EC). Copernicus Open Access Hub. Available online: https://scihub.copernicus.eu/dhus/#/home (accessed on 17 Nov. 2018).
133. Centre national d'études spatiales (CNES). CNES website. Available online: https://theia.cnes.fr/atdistrib/rocket/#/search?collection=SENTINEL2 (accessed on 17 Nov. 2018).
134. Quinlan, P. Marr's Vision 30 years on: From a personal point of view. *Perception* **2012**, *41*, 1009–1012.
135. Poggio, T. *The Levels of Understanding Framework*; Technical Report, MIT-CSAIL-TR-2012-014, CBCL-308; Computer Science and Artificial Intelligence Laboratory: Cambridge, MA, USA, 31 May 2012.
136. Lu, D.; Weng, Q. A survey of image classification methods and techniques for improving classification performance. *Int. J. Remote Sens.* **2007**, *28*, 823–870.
137. Ghosh, D.; Kaabouch, N. A Survey on Remote Sensing Scene Classification Algorithms. *WSEAS Trans. Signal Proc.* **2014**, *10*, 504–519.
138. Cheng, G.; Han, J. A survey on object detection in optical remote sensing images. *ISPRS J. Photogramm. Remote Sens.* **2016**, *117*, 11–28.
139. Zhou, G. *Architecture of Future Intelligent Earth Observing Satellites (FIEOS) in 2010 and Beyond*; Technical Report; National Aeronautics and Space Administration Institute of Advanced Concepts (NASA-NIAC): Washington, DC, USA, 2001.
140. GISCafe News. Earth-i Led Consortium Secures Grant from UK Space Agency. 19 July 2018. Available online: https://www10.giscafe.com/nbc/articles/view_article.php?section=CorpNews&articleid=1600936 (accessed on 15 July 2018).
141. Planet Labs. Planet Surface Reflectance Product. 2018. Available online: https://assets.planet.com/marketing/PDF/Planet_Surface_Reflectance_Technical_White_Paper.pdf (accessed on 11 July 2018).
142. Baraldi, A. Impact of radiometric calibration and specifications of spaceborne optical imaging sensors on the development of operational automatic remote sensing image understanding systems. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2009**, *2*, 104–134.
143. Pacifici, F. Atmospheric Compensation in Satellite Imagery. U.S. Patent 9396528B2, 19 July 2016.
144. Pacifici, F.; Longbotham, N.; Emery, W.J. The Importance of Physical Quantities for the Analysis of Multitemporal and Multiangular Optical Very High Spatial Resolution Images. *IEEE Trans. Geosci. Remote Sens.* **2014**, doi:10.1109/TGRS.2013.2295819.
145. Piaget, J. *Genetic Epistemology*; Columbia University Press: New York, NY, USA, 1970.