The Importance of Physical Quantities for the Analysis of Multitemporal and Multiangular Optical Very High Spatial Resolution Images

Fabio Pacifici, Senior Member, IEEE, Nathan Longbotham, Member, IEEE, and William J. Emery, Fellow, IEEE

Abstract—The analysis of multitemporal very high spatial resolution imagery is too often limited to the sole use of pixel digital numbers which do not accurately describe the observed targets between the various collections due to the effects of changing illumination, viewing geometries, and atmospheric conditions. This paper demonstrates both qualitatively and quantitatively that not only physically based quantities are necessary to consistently and efficiently analyze these data sets but also the angular information of the acquisitions should not be neglected as it can provide unique features on the scenes being analyzed. The data set used is composed of 21 images acquired between 2002 and 2009 by QuickBird over the city of Denver, Colorado. The images were collected near the downtown area and include single family houses, skyscrapers, apartment complexes, industrial buildings, roads/highways, urban parks, and bodies of water. Experiments show that atmospheric and geometric properties of the acquisitions substantially affect the pixel values and, more specifically, that the raw counts are significantly correlated to the atmospheric visibility. Results of a 22-class urban land cover experiment show that an improvement of 0.374 in terms of Kappa coefficient can be achieved over the base case of raw pixels when surface reflectance values are combined to the angular decomposition of the time series.

Index Terms—Angular decomposition, bidirectional reflectance distribution function (BRDF), multiangular analysis, multitemporal analysis, optical very high spatial resolution, QuickBird (QB), surface anisotropy, surface reflectance, urban change detection, urban classification, WorldView-2.

I. INTRODUCTION

THE commercial availability of optical very high spatial resolution spaceborne imagery began more than ten years ago with the launch of IKONOS and QuickBird (QB), which led to an increasing interest in satellite data for mapping and precise location-based service applications. Since then, a large amount of data has been acquired, including images from newer and more complex platforms such as WorldView-1 and WorldView-2, GeoEye-1, and the more recent Pléiades-1A and Pléiades-1B. Currently, the global capacity of the very high spatial resolution imaging satellites is greater than 1.8 billion square kilometers per year (which corresponds to more than 12 times the land surface area of the Earth) and is expected to increase to more than 2.4 billion square kilometers per year (about 16 times the land surface area of the Earth) in the near future. Much of this imagery is collected with a wide range of azimuth and elevation angles. Fig. 1 shows a typical 30-day collection by the QB, WorldView-1, and WorldView-2 constellation.

The availability of submeter resolution data regularly acquired over the same geographical region has proved to be effective in developing various monitoring systems, from precision agriculture (including the growing and harvesting of crops) to disaster management and search and rescue operations in the case of natural events (earthquakes, hurricanes, and floods). In urban areas, multitemporal data provide information about newly built constructions or the demolition of existing structures, road conditions, and urban growth. Furthermore, the WorldView class of satellites has a high-performance camera control system capable of rapid retargeting, allowing collection within a few seconds of dozens of images over a single target, each with a unique angular perspective. This capability opens a unique approach to multitemporal imaging whose applications have been discussed in [1].

Despite the large amount of data acquired and the progress of space technology in designing and launching more sophisticated sensors, very little research addresses the advantages and challenges of multitemporal optical very high spatial resolution data. This has been confirmed by a very recent special issue on the analysis of multitemporal remote sensing data [2] where only one contribution dealt with submeter optical imagery [3].

In the remote sensing literature, two opposite approaches are generally considered to analyze the data sequence in the
context of land cover mapping: A set of independent models is developed for each image, or a unique model is generated from the entire set of images at once. The first approach does not guarantee that the temporal information is fully exploited as each classifier is tailored to fit a specific subproblem. On the other hand, a generalized model may suffer from different data distributions in the image sequence, resulting in poor results. In [4], Heas and Datcu propose an unsupervised method to learn trajectories of dynamic clusters followed by an interactive learning process. The trajectories in the feature space result in graphs coding spatiotemporal structures contained in the data sequence. The information-bottleneck principle is introduced in [5], which combines model selection with rate-distortion analysis in order to determine the optimal number of clusters. In [6], Petitjean et al. propose an approach to deal with irregularly sampled time series based on the dynamic-time-warping concept, while various similarity measurements are used to model consecutive image pairs in [7].

Very recently, there has been a large number of publications in the remote sensing literature on domain adaptation, whose goal is to adapt a prediction function from a source domain to a target domain, reducing the effects of shifts between different, but related, data sets (such as one gathered from multitemporal acquisitions). In [8], Bruzzone and Prieto exploit the distribution of a new image to reestimate the parameters of a maximum likelihood classifier. In [9], a binary hierarchical classifier is used in the target domain to leverage the information extracted from the existing labeled data. In [10], the support vectors of a support vector machine classifier are iteratively adapted to the distribution of a new domain. Active learning methods have also been considered in [12]–[14] to cope with data set shifts. While most of the domain adaptation methods deal with adjusting the model to the target domain, Tuia et al. propose a method described in [15] where data manifolds are deformed through nonlinear transformations driven by a graph matching procedure aimed at finding correspondences between domains, whereas Leiva-Murillo et al. introduce in [16] the concept of multitask learning by jointly solving a set of prediction problems by sharing information across different tasks.

In most remote sensing studies of optical very high spatial resolution imagery, however, the analysis of time series is limited to the use of pixel digital numbers (DNs), ignoring the physical effects of atmospheric, viewing, and illumination changes between image collections. The importance of radiometric calibration (and the need to work with physical quantities) has already been suggested in [17] for the implementation of operational automated remote sensing image understanding systems. As previously discussed in [18], results over the past few years have offered very modest improvements with respect to one obtained from other methods (usually, less than 1%–2% in absolute terms) on a limited number of classes, which frequently only includes conventional targets, such as man-made structures, vegetation, soil, and water, and number of images (two to three, in general). Often, the various results are not even discussed in terms of their statistical significance, leaving the reader wondering if the principles of the proposed technique are repeatable on different data sets or if the improvements were rather obtained by strenuously exercising the data samples to achieve the desired output.

Another limitation of current techniques is that multitemporal data sets are generally analyzed only considering the temporal domain. Instead, the temporal information should be coupled to the corresponding angular component to make the best use of the available imagery. In fact, the radiometric differences in time series may often be corrected or accounted for by understanding the physics of the acquisitions. The objective of this paper is not only to demonstrate that physical quantities are necessary to consistently and efficiently analyze submeter optical imagery but also to bring attention to the research community that the angular information of the acquisitions should not be neglected as unique features can be derived from it. To the best knowledge of these authors, this is the first time that these aspects are addressed in the context of land cover mapping and change detection of urban areas using optical very high spatial resolution image time series.

The data set used is composed of 21 images acquired between 2002 and 2009 by QB over the city of Denver, Colorado. The time series covers part of the downtown area and includes single family houses, skyscrapers, apartment complexes, industrial buildings, roads/highways, urban parks, and bodies of water. The acquisition dates are reported in Table I, and the tempo–angular distribution (in zenith and azimuth angles) of the image sequence is shown in Fig. 2 along with the Sun...
position for the specific day of the year (also reported in Table I). All images were acquired within a 30° zenith angle and fairly evenly distributed azimuth angles, while the Sun position exhibits the natural declination through the year (being closer to 20° in zenith during summers and to 70° during winters). All 21 images were converted to surface reflectance by an automatic DigitalGlobe proprietary method designed for very high spatial resolution panchromatic or visible/near-infrared imagery [19] (whose performances are discussed in detail in [20]).

The rest of this paper is organized as follows. In Section II, the differences between nonphysical and physical quantities are described in detail, including the theoretical formulation to retrieve at-sensor radiance, at-sensor reflectance, and surface reflectance, while the effects of surface anisotropy and two different angular decomposition models are reviewed in Section III. Both qualitative and quantitative experimental results are discussed in Section IV to illustrate the advantages of physical quantities and angular decompositions for the analysis of multitemporal data sets. Discussion and conclusions are summarized in Sections V and VI, respectively.

II. Nonphysical and Physical Quantities

The accurate analysis of multitemporal remote sensing data depends upon the ability to distinguish between relevant and nonrelevant changes on the Earth surface through time [21]. At the same time, the capability to detect and quantify these changes depends on consistent sensor measurements, which are generally distorted by changing atmospheric conditions, solar illumination, and satellite viewing geometries. Therefore, it is preferable to convert the raw image counts to physical quantities that are capable of accurately describing the imaged surfaces before analyzing the data from a single scene, between images acquired on different dates, or by different sensors.

With reference to Fig. 3, optical remote sensing satellites measure photoelectric signals that are equivalent to the radiance reflected by the Earth surface when illuminated by the Sun and perturbed by the atmosphere, including the effect of gaseous absorption and scattering by molecules and aerosols. These measured signals are not directly accessible to the end users as they are converted and stored as DNs.

The DN values [counts] are proportional to the radiance \( L [W m^{-2} sr^{-1} \mu m^{-1}] \) entering the telescope aperture according to [22]

\[
DN = L \cdot GAIN + OFFSET
\]

where \( GAIN \) is the absolute gain \([counts/W m^{-2} sr^{-1} \mu m^{-1}]\) and \( OFFSET \) is the instrument offset \([counts]\). This formulation assumes that the detectors have a linear response as a function of input radiance. The appropriate gain and offset settings are wavelength dependent and are operationally selected based on the following:

1) compression levels;
2) pixel aggregation;
3) line rate;
4) bit depth;
5) time delay integration (TDI), which increases the exposure provided by the basic line rate;
6) seasonality (the TDI setting for a given image is selected based on the estimated solar elevation angle).

This means that image DN counts are unique not only to the sensor but also to the very specific operational setting selected for the acquisition. Additionally, an ideal object with a spectral signature that is invariant over time may show significant DN differences in multitemporal data sets acquired with similar operational settings due to different atmospheric conditions and/or different viewing and solar geometries. In this sense, DN data, which do not represent physical quantities, should not be directly compared to DN imagery from other sensors nor even between images from the same sensor as collection settings, atmospheric effects, and viewing and illumination geometries may be significantly different.

The calculation of at-sensor, also known as top-of-atmosphere (TOA), radiance is a necessary step for converting the image into a physically meaningful common scale [23]. Even though end users have access only to the image DN counts, pixel values can easily be converted to at-sensor radiance by inverting (1) and using the \( GAIN \) and \( OFFSET \) information provided in the image metadata.

An additional reduction in scene-to-scene variability can be achieved by converting the at-sensor radiance \( L \) to TOA reflectance \( \rho^{TOA[unitless]} \) using

\[
\rho^{TOA} = \frac{L \cdot d_{ES}^2 \cdot \pi}{E_{sun} \cdot \cos(\theta_S)} \tag{2}
\]

where \( E_{sun} \) is the mean exoatmospheric solar irradiance \( [W m^{-2} \mu m^{-1}] \), \( \theta_S \) is the solar zenith angle [degrees], and \( d_{ES} \) is the Earth–Sun distance [astronomical units] as derived in the Appendix. There are several benefits for using TOA reflectance with respect to TOA radiance, such as the removal of the cosine effect of different solar zenith angles, the compensation for different values of the exoatmospheric solar irradiance arising from spectral band differences, and the correction for the variation in the Earth–Sun distance between the different acquisitions [23].

To estimate surface reflectance \( \rho[unitless] \) from satellite data, TOA radiance needs to be compensated for atmospheric absorption and scattering phenomena, approximating what
would be measured by a sensor held just above the Earth surface, without any alterations from the atmosphere [24]. One of the advantages of surface reflectance is the physically based normalization of the image values throughout the dates regardless of the specific material, can show different properties of the surface material [29]. In general, both natural and man-made surfaces show some degree of spectral variability due to Rayleigh scattering, ranging from about 0.15 to 0.21 at 427 nm (to be compared with both the reference hyperspectral and the retrieved surface reflectance signatures that range from about 0.02 to 0.04). On the other hand, TOA reflectance values at 949 nm are smaller than both the reference hyperspectral and the retrieved surface reflectance signatures due to water vapor absorption. It is important to account for these distortions. For example, the average of the normalized difference vegetation index calculated from TOA reflectance values is about 0.688, to be compared to 0.814 in case of surface reflectance, which corresponds to a difference of more than 15%. Therefore, surface reflectance not only provides consistent quantities across the various bands at different times, geographical locations, and atmospheric conditions but also minimizes the spectral distortions with respect to the reference hyperspectral signatures, making it a suitable transformation for the analysis of large temporal data sets.

A comprehensive review of remote sensing quantities can be found in [28].

III. ACCOUNTING FOR ANGULAR VARIABILITY

In the previous section, it was mentioned that the Sun–Earth–satellite geometry influences the measured at-sensor radiance and, therefore, both DN and surface reflectance values. While the Earth–satellite geometry can be kept constant with nadir-looking sensors such as the Landsat series, commercial satellites operate with significantly varying geometries due to their agility in the rapid retargeting of the camera. This is especially important in the case of emergencies when the viewing geometries are opened up to maximize the collection opportunities.

In this section, geometrical effects are discussed in terms of surface anisotropy whose main mechanisms are illustrated in Section III-A. In Section III-B, two angular decomposition models are reviewed in detail.

A. Surface Anisotropy

A surface that reflects the incident energy equally in all directions is said to be Lambertian, and its reflectance is invariant with respect to illumination and viewing conditions. On the contrary, a surface is said to be anisotropic when its reflectance varies with respect to illumination or viewing geometries. These changes are driven by the optical and structural properties of the surface material [29]. In general, both natural and man-made surfaces show some degree of spectral
anisotropy, and this behavior can be described by the bidirectional reflectance distribution function (BRDF) [30], [31].

The fundamental components of surface anisotropy as described in [32] are illustrated in Fig. 5:

1) surface scattering, which can be observed when forward scattering elements are present and includes specular reflection;
2) radiative transfer-type volumetric scattering, which is due to the presence of finite scatterers, such as the leaves of plants;
3) geometrical–optical scattering, which is produced by casting shadow and mutual obscuration of vertical surfaces.

Consequently, multiangular acquisitions contain information about the physical structure and characteristics of the observed target. Depending on the study, surface anisotropy can be seen as a source of noise (e.g., when analyzing spectral signatures of time series) or, alternatively, as a source of information in addition to the tempo–spectral dimensions [30].

To qualitatively show the effects of surface anisotropy across multitemporal very high spatial resolution data sets, two examples are discussed as follows.

1) The two images in Fig. 6 were acquired over Denver only five days apart in March 2008, but from opposite satellite azimuths, i.e., 150° and 339° (see Table I for details). As shown, some surfaces appear brighter when observed from the backward direction, while others appear brighter when observed from the forward direction. Therefore, image matching techniques that do not account for localized distributions will necessarily fail, possibly compromising the spectral meaning of the data.

2) Fig. 7 illustrates two domain representations of the Denver baseball stadium grass during only the summer acquisitions of the time series. Fall or winter dates were not considered to avoid spectral variations related to the state of grass rather than to the different viewing conditions.

This choice also minimizes the Sun declination, assuring a quasi-consistent source of illumination, in terms of position and intensity. Finally, it is worth mentioning that the grass of the stadium is kept uniform through the season according to baseball regulations. This means that this data set is suitable to illustrate the effects of different viewing conditions when illumination and target are unchanged. Specifically, the multitemporal plot in Fig. 7(a) illustrates variations across the dates, with the near-infrared band ranging from about 0.44 to 0.62 just in the summer of 2005. The polar plot in Fig. 7(b) shows that these variations are consistent with the angular properties of the surface and the different viewing geometries (with elements of the previously discussed scattering mechanisms). In fact, observations in the backward direction consistently appear brighter than that in the forward direction.

B. Angular Decomposition Models

The increasing availability of multiangular measurements from spaceborne sensors, such as the Along-Track Scanning Radiometer-2, Polarization and Directionality of the
Su Chopping et al. shown by Kimes improve the classification accuracy of various land covers as medium and high spatial resolution angular data significantly anisotropy. Several studies have proven over the years that new and unique opportunities to understand and exploit surface WorldView-1, and WorldView-2, or aerial cameras has brought (MISR), Compact High Resolution Imaging Spectrometer, Earth’s Reflectances, Multiangle Imaging Spectro-Radiometer direction. The backward direction consistently appear brighter than that in the forward of the surface and the different viewing geometries. In fact, observations in (b) shows that these variations are consistent with the angular properties in (b) near-infrared angular reflectance (with nonconstant illumination during only the summer acquisitions of the time series: (a) multitemporal sig-

Fig. 7. Two domain representations of the Denver baseball stadium grass during only the summer acquisitions of the time series: (a) multitemporal sign-

Earth’s Reflectances, Multiangle Imaging Spectro-Radiometer (MISR), Compact High Resolution Imaging Spectrometer, WorldView-1, and WorldView-2, or aerial cameras has brought new and unique opportunities to understand and exploit surface anisotropy. Several studies have proven over the years that medium and high spatial resolution angular data significantly improve the classification accuracy of various land covers as shown by Kimes et al. [33], Sandmeier and Deering [34], Chopping et al. [35], and Armston et al. [36]. More recently, Su et al. [31], Verrelst et al. [37], Laurent et al. [38], and Koukal and Atzberger [39] showed that angular data are useful for vegetation mapping as they provide information that is not available in the spectral domain. Some studies addressing multiaxial observations were recently reviewed in a dedicated special issue [1].

A number of techniques have been proposed to characterize and efficiently use angular information. Two assumptions generally hold true [40].

1) Atmospheric effects should be removed.
2) A sufficient angular sampling should be available to provide robust retrievals.

In particular, Koukal and Atzberger [39] found that angular observations should cover both the backward and forward scattering directions. Otherwise, models will overfit the data, and the parameters retrieved will not represent the true anisotropy of the target.

Angular models can be classified as physical or empiri-
cal [41]. Physical models rely on first-principle physics and require a complete and comprehensive model parameterization (e.g., inputs such as surface roughness or complex refractive index), which is often very difficult to obtain. On the contrary, empirical models rely exclusively on measured angular values. A tradeoff between these two techniques is represented by semiempirical models which incorporate measured data to elements of physics-based principles. Semiempirical models can be applied without any knowledge of the target complexity and composition, as they do not impose severe hypotheses about the nature and structure of the surface being modeled [29]. Among the various semiempirical models proposed, the kernel-driven Ross–Li [42] and the Rahman–Pinty–Verstraete (known as RPV) [29] are some of the most widely used. The former assumes that surface anisotropy can be described by the linear contribution of a set of kernels that describe the basic mechanisms of surface anisotropy, whereas the latter provides a representation of surface anisotropy by means of angular functions [30]. In particular, the Ross–Li model is the basis for the MODIS BRDF/albedo product, while the the RPV model is used to generate the MISR BRDF/albedo product. It was found in [43] and [44] that both models provide comparable results, with errors within 10% from observations [45].

Given the solar zenith angle $\theta_s$, the view zenith angle $\theta_v$, and the relative view-solar azimuth angle $\phi$, the Ross–Li model decomposes the observed angular surface reflectance into three basic scatter mechanisms: isotropic, radiative transfer-type volumetric, and geometrical–optical. This model combines these elements as

\[
\rho_\text{Ross–Li}(\theta_s, \theta_v, \phi) = f_{\text{iso}} + f_{\text{vol}} \cdot K_{\text{vol}}(\theta_s, \theta_v, \phi) + f_{\text{geo}} \cdot K_{\text{geo}}(\theta_s, \theta_v, \phi)
\]  

(4)

where $K_{\text{vol}}$ and $K_{\text{geo}}$ are the volumetric and geometric scattering kernels and $f_{\text{iso}}, f_{\text{vol}},$ and $f_{\text{geo}}$ are the isotropic, volumetric, and geometrical kernel scaling factors. The isotropic scattering has no dependence on the incidence or viewing angle and therefore does not have a geometrically dependent kernel. The angular behavior of the volumetric kernel presents a minimum near the backward direction and bright limbs, while the angular behavior of the geometrical kernel shows a maximum in the backward direction, where there are no shadows [40].
The RPV model decomposes the observed angular surface reflectance into three independent components, representing the amplitude $\rho_0$, the shape anisotropy $k$, and the asymmetry factor $\Theta$, according to

$$\rho_{\text{RPV}}(\theta_s, \theta_v, \phi) = \rho_0 \cdot \frac{\cos^{k-1} \theta_s \cos^{k-1} \theta_v}{(\cos \theta_s + \cos \theta_v)^{1-k}} \cdot G(g, \Theta) \cdot H(G) \cdot F(g, \Theta)$$

with

$$F(g, \Theta) = \frac{1 - \Theta^2}{(1 + \Theta^2)^{1/2}}$$

$$H(g, \rho_0) = 1 + \frac{1 - \rho_0}{1 + G}$$

$$\cos g = \cos \theta_s \cos \theta_v + \sin \theta_s \sin \theta_v \cos \phi$$

$$G(\theta_s, \theta_v, \phi) = (\tan^2 \theta_s + \tan^2 \theta_v - 2 \tan \theta_s \tan \theta_v \cos \phi)^{1/2}$$

The parameter $\rho_0 \in [0, 1]$ characterizes the intensity of the target, but it should not be confused with the single-scattering albedo or the true reflectance of the target, as it is independent of the angular variations. The parameter $k \in [0, 2]$ indicates the anisotropy of the target. Values of $k$ smaller than 1.0 represent a bowl-shape anisotropy pattern, where $k$ increases with the view zenith angle. In contrast, values of $k$ greater than 1.0 represent a bell-shape anisotropy pattern, where $k$ reaches its maximum at the nadiral view. A Lambertian surface is represented by the ideal case of $k = 0$. It is important to emphasize that $k$ is influenced by the direction of the illumination with respect to the target. Therefore, the values of $k$ in multitemporal data sets should be carefully interpreted not as an intrinsic property of the surface, i.e., $k$ may vary as a function of the season [30]. Finally, the asymmetry factor $\Theta \in [-1, 1]$ controls the relative amount of forward, $\Theta \in (0, 1]$, and backward scattering, $\Theta \in [-1, 0)$.

A more applicable discussion about the retrieval of the RPV model (the triplet $\rho_0$, $k$, and $\Theta$) is discussed in Section IV-D. The interested reader can also refer to [39].

IV. Experimental Results

In the previous sections, it was discussed how surface reflectance provides a consistent feature space and that an additional dimensionality can be exploited by including angular information. The results of four different experiments carried out over the Denver time series are described in this section to support these concepts. The first two exercises focus only on the temporal aspects of the data set and on the advantages of working in the surface reflectance domain. In particular, Section IV-A illustrates the analysis of the tempo–spectral variations of a rooftop, while Section IV-B addresses the differences of two automated urban change detection methods over two sets of raw DNs and surface reflectance image pairs. The other two exercises discuss the advantages of coupling the temporal and the angular information: In Section IV-C, the tempo–angular spectral signatures of an “unknown object” are analyzed with the aim of providing additional information about its characteristics, whereas Section IV-D provides the results of a 22-class urban land cover exercise.

A. Analysis of Multitemporal Spectral Signatures

The multitemporal spectral signatures of a flat roof are shown in Fig. 8 in both DNs (normalized to 1.0 for the sake of comparison) and surface reflectance values. These spectral signatures represent the average of all pixels over the roof. The aerosol optical depth (AOD), which is inversely related to the visibility (i.e., the lower the AOD, the higher the visibility), is also reported for completeness. The AOD is one of the outputs provided by the method described in [19] and [20].

The DN curves in Fig. 8(a) show that the green values are consistently higher than the other spectral components, with very large variability in all bands through the years. This leads to the conclusion that the color of the roof is consistent with some shade of green, and no additional information can be reliably derived to explain, for example, the nature of the temporal behavior.

On the other hand, three well-defined fairly stable temporal regions can be identified from the surface reflectance values shown in Fig. 8(b). Specifically, there is a relatively flat spectral plateau at about 0.35 reflectance, followed by a much darker response between the winter of 2007 and the spring of 2008 and, then, by a highly reflecting region in subsequent periods to the end of the time series (about 0.70 reflectance). These relatively flat temporal regions, and their sharp transitions, may indicate that the roof being investigated went through renovation during the years. Also, there is a small difference between the bands at any date in the sequence (about 0.05 or less in absolute terms), suggesting that the roof is consistent with three different shades of gray, from a medium tone up to 2007 to a very bright shade after 2008, with a dark response in between.

The temporal behavior can be understood looking at Fig. 8(c), which shows that the roof was being remodeled between the end of 2007 and the beginning of 2008. Therefore, a simple analysis of surface reflectance signatures indicates that a physical input space can be extremely helpful in understanding the temporal variability and its sharp transitions, whereas a DN-based analysis failed to provide consistent information. It is also worth mentioning that the date-to-date variability of the DN temporal signatures in Fig. 8(a) is related to the AOD values with a Pearson correlation coefficient of 0.427, which is significant at the 0.05 confidence level. However, the date-to-date variability of the surface reflectance temporal signatures in Fig. 8(b) is correlated with the AOD values with a Pearson coefficient of 0.097, which is no longer significant at the 0.05 level. This emphasizes the fact that atmospheric conditions significantly affect the DN input domain, whereas the date-to-date variability of the surface reflectance curves can be attributed to the anisotropic nature of the target. This is further discussed in Section IV-C, which illustrates the analysis of multitemporal spectral signatures coupled to their angular components.

B. Automated Urban Change Detection

Several methodologies have been developed during the years in the context of automated change detection. Two widely used approaches are based on change vector analysis (CVA)
and principal component analysis (PCA). In the former, pixels are represented by their vectors in the feature space, and the changes are derived as the difference of the feature vectors between the images [46]–[49]. In the latter, the first principal component (which corresponds to the largest eigenvalue) reflects the unchanged parts of a set of images, whereas changes can be depicted from the components corresponding to smaller eigenvalues [50]–[53].

Two sets of raw DNs and surface reflectance image pairs are used to qualitatively and quantitatively investigate the differences of nonphysical and physical domains for change detection studies. The image pairs correspond to a subset of the entire scene where several changes occurred between July 2002 and August 2008 (shown in Fig. 9(a) and (b), respectively). These changes include the construction and demolition of several large buildings in the industrial area and the remodeling of various roofs in the residential district. The two dates were selected as the corresponding images were acquired with similar viewing geometries and during similar time of year, minimizing both stereoscopics and seasonal effects (which are not the focus of this analysis). The near-infrared band was not considered to filter out changes in vegetation cover, which are considered not relevant in this analysis.

Fig. 9 shows the CVA and PCA change detection results derived using DN counts as input (see Fig. 9(c) and (e), respectively) and using surface reflectance values (see Fig. 9(d) and (f), respectively), where the magnitude of change is normalized to the interval \([-1.0, +1.0]\). Extensive false alarms (FAs) are visible, particularly over paved surfaces and rooftops, in both change maps that were produced from DN counts. On the other hand, the change maps produced from surface reflectance values clearly identify the changes due to the construction of new buildings (shown in cold colors) and the demolition of older structures (shown in warm colors).

Despite the simplicity of both methodologies, the improvement obtained from the use of surface reflectance data is evident. This is quantitatively illustrated in Table II which reports the overall accuracy (OA), FAs, and missed alarms (MAs) for the four cases and six threshold levels. As expected, the larger the threshold, the smaller the FA and the larger the MA rates. The goal of automated, semiautomated, and manual thresholding techniques is to find a value that maximizes the OA and minimizes both the FA and the MA. For example, small thresholds can be used when there is no bias between the average values of the two acquisitions, whereas larger thresholds are necessary to account for differences in the data distributions. Table II shows that low thresholds (i.e., 0.05 and 0.10) provide high accuracy with low FA and MA values only for the surface reflectance cases. It is interesting to point out that, for both CVA-DN and PCA-DN, only a large threshold (0.25 or 0.30) can result in an acceptable FA rate, but at the price of a MA rate well above 30%.

C. Combined Analysis of Multitemporal and Multiangular Spectral Signatures

Similar to Fig. 8(a), the AOD and the multitemporal spectral signature in normalized DNs of an “unknown object” are
Fig. 9. Two scenes used to generate the change detection maps were acquired on (a) July 2002 and (b) August 2008. The dates were selected as the corresponding images were acquired with similar viewing geometries and during similar time of year, minimizing both stereoscopics and seasonal effects (which are not the focus of this analysis). The change detection maps in (c) and (e) were derived from DN counts as input using the CVA and PCA approaches, respectively, whereas the change maps in (d) and (f) were derived from surface reflectance values using the CVA and PCA approaches, respectively. Extensive FAs are visible, particularly over paved surfaces and rooftops, in both change maps that were produced from DN counts. On the other hand, the change maps produced from surface reflectance values clearly identify the changes due to the construction of new buildings (shown in cold colors) and the demolition of older structures (shown in warm colors).
illustrated in Fig. 10(a). With this plot, it is very difficult to extract useful information on the nature of the surface being investigated. For example, there are evident variations in the near-infrared band between winter and summer acquisitions (from about 0.16 to 0.36), which may be indicative of a natural surface (such as vegetation). However, it is not possible to guess with confidence the color of the object as the DN values are influenced by several factors as discussed previously. In this particular case, one can believe that the object may be yellow, as the red and green bands are very close in amplitude to each other. Also, in this case, it is worth mentioning that the date-to-date variability is related to the AOD values with a Pearson coefficient of 0.538, which is significant at the 0.05 level. On the other hand, from Fig. 10(b), which represents the same plot as in Fig. 10(a), but in the surface reflectance domain, it is possible to deduce that the “unknown object” has a smaller seasonal variability in the near-infrared band (from 0.26 to 0.35) with peaks not necessarily corresponding to the warmer months (see, for example, the image acquired on January 2009), which can lead to the conclusion that the object may not be a natural surface. Furthermore, the “unknown object” is red, as the green and the blue spectral signatures are constantly below the red one. Therefore, using only a multitemporal analysis, it is possible to assess (with some degree of uncertainty) that the red object, possibly man-made. In this case, the date-to-date variability is consistent with the assumptions of it is possible to assess (with some degree of uncertainty) that the object may not be a red object, possibly man-made. In this case, the date-to-date variability is consistent with the assumptions of red object, possibly man-made.

Another example of angular surface reflectance for a nonconstant illumination geometry is reported in Fig. 10(d), which represents a white roof, as shown in the background image. In particular, the angular surface reflectance shows a fairly consistent plateau of about 0.35 in the azimuthal southern part of the plot, and two areas with much higher reflectance (about 0.60) in the Sun specular reflection. From this, it is possible to understand that the roof is also pitched (because of the two disjoint areas of higher reflectance), and the material is most likely metal (because of the high response in the specular direction).

It is worth mentioning that, as for Section IV-A, the analysis of the spectral signatures represents the average of all pixels over each of the two roofs.

D. Urban Land Cover Classification

In order to quantitatively investigate the benefits of surface reflectance and angular decompositions to improve the urban land cover classification of image time series, 22 noncommon classes of interest were selected from the Denver data set. These classes include different kinds of grass, water, soil, paved surfaces, and pitched or flat roofs as shown in Table III. The angular decomposition used for this experiment is RPV [30].

Three representation domains of the time-series data set, DN, surface reflectance, and RPV, were used in three independent classification experiments. The 21 images were randomly sampled by 101 different cross-validation runs, where 11 dates were retained for training and the remaining ones were retained for validation. These independent sets were used to create two distinct RPV models (one for training and one for validation). A random forest model with 100 trees was generated for each cross-validation run and each input domain [26]. It is worth noting that the models for the DN and surface reflectance domains assumed an input space composed by four features, as the RPV decomposition provides \( \rho_0, k, \) and \( \Theta \) for each spectral band.

The RPV decomposition was implemented by fitting the model to different lookup tables (LUTs), one for each acquisition, as the illumination and viewing geometries change for each date. In particular, the RPV parameters were sampled for the generation of the LUTs according to the following scheme.

1) \( \rho_0 \in [0.0 : 0.01 : 1.0] \),
2) \( k \in [0.0 : 0.01 : 2.0] \),
3) \( \Theta \in [-1.0 : 0.01 : 1.0] \).
Successively, for each class of interest, the 11 LUT entries that provided the smallest root mean squared error between the retrieved values and the one measured by the sensor were retained, and their average value was considered as the solution of $\rho_0$, $k$, and $\Theta$.

Table III

<table>
<thead>
<tr>
<th>Classes of Interest</th>
<th>natural surfaces</th>
<th>man-made surfaces</th>
</tr>
</thead>
<tbody>
<tr>
<td>grass type 1 (stadium)</td>
<td>parking lot type 1 (asphalt, bright)</td>
<td>parking lot type 2 (asphalt, dark)</td>
</tr>
<tr>
<td>grass type 2 (golf-fairway)</td>
<td>roof type 1 (brown, dark, flat)</td>
<td>roof type 2 (brown, bright, flat)</td>
</tr>
<tr>
<td>grass type 3 (golf-green)</td>
<td>roof type 3 (concrete, gray, flat)</td>
<td>roof type 4 (concrete, bright, flat)</td>
</tr>
<tr>
<td>grass type 4 (park 1)</td>
<td>roof type 5 (concrete, dark, flat)</td>
<td>roof type 6 (red, brick, pitched)</td>
</tr>
<tr>
<td>grass type 5 (park 2)</td>
<td>roof type 7 (blue, pitched)</td>
<td>roof type 8 (metal, pitched)</td>
</tr>
<tr>
<td>tree</td>
<td>shadow (of various man-made materials)</td>
<td></td>
</tr>
<tr>
<td>water type 1 (lake)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>water type 2 (river)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>soil type 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>soil type 2 (play-ground)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>soil type 3 (baseball field)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig. 10. Multitemporal spectral signatures in (a) normalized DNs and (b) surface reflectance for the pitched brick-red roof shown in background in (c) which illustrates its angular surface reflectance of the red band (with natural nonconstant illumination). The yellow circles represent the Sun locations throughout the time series in zenith and azimuth angles, whereas the black dots correspond to the satellite positions. An additional angular plot for a metallic pitched roof (shown in background) is illustrated in (d).

Fig. 11 illustrates the retrieved near-infrared values of $\rho_0$, $k$, and $\Theta$ for the five types of grass. For the sake of completeness, the near-infrared values in DNs and surface reflectance domains are also reported. As shown, it is not possible to accurately define more than one single cluster of grass with DN values due to the large degree of overlap between the five distributions. On the other hand, the surface reflectance domain reduces the intraclass variability of the different types, allowing the identification of two separate clusters of healthy and less healthy grass (grass type 1-2-3 and grass type 4-5, respectively) with a threshold of 0.48. Finally, the triplet $\rho_0$, $k$, and $\Theta$ allows the discrimination of all five grass types with fairly good confidence (as discussed later in more detail). In particular, the intensity $\rho_0$ contains similar information as that provided by the surface reflectance domain, making it possible to cluster healthy and less healthy grass with a threshold of 0.31. The anisotropy factor $k$ has values smaller than 1.0 for all grass types but grass type 1, representing a bowl-shape pattern as found in [54]. The class grass type 1 shows an anisotropy factor...
consistent with a bell-shape pattern which can be explained by the renovations of the stadium field during winter acquisitions, when the grassy surface is generally converted to bare soil [55]. This also explains the larger variability of this class with respect to the others in all domains. Furthermore, as discussed earlier, the anisotropy factor \( k \) is more sensitive to temporal changes [30] as shown by the larger intraclass variability compared to \( \rho_0 \) and \( \Theta \). Overall, \( k \) differentiates grass type 1 from grass type 4 and the other grass types (with thresholds 1.0 and 0.8, respectively). The asymmetry factor \( \Theta \) is systematically below 0.0 for all types of grass, which is consistent with the backward scattering behavior as also evinced from Fig. 7. The use of \( \Theta \) allows the discrimination of grass type 2 from grass type 3 (with a threshold of \(-0.09\)).

Fig. 12 illustrates the classification accuracies (in terms of Kappa coefficient) for the three domains. The box plots account for the variability of both the different random forest model initializations and training/testing sets of the 101 cross-validation runs. The three results are statistically significant according to the McNemar test. The average accuracy when DN values are used as input is about 0.477. This result is quite in line with initial expectations as some of the classes are spectrally similar to each other, such as the five types of grass, or they may require knowledge of the structure being analyzed, such as the pitched roof classes. The \( F_1 \) measure of the various targets is reported in Table IV. As shown, the grass classes are not differentiated, with a maximum \( F_1 \) just slightly above 0.3. In general, the remaining classes have an \( F_1 \) value around 0.5. However, it is interesting to note how roof type 6 and roof type 7 are actually quite well discriminated with an \( F_1 \) value of over 0.85. This may be due to the very distinct spectral components of these two roofs (which appeared red and blue in the visible bands, respectively) with respect to the other classes. An improvement over the DN case of about 49.1% is achieved by considering surface reflectance values, corresponding to a Kappa coefficient of 0.711. Roughly all classes show better accuracies even though the grass and soil clusters
are still not well differentiated, which can be explained by the spectral similarity of these targets. Finally, a Kappa coefficient of 0.851 is achieved when surface reflectance is combined to the RPV decomposition of the time series, corresponding to an improvement of 78.4% over the base case of DNs. The \( F_1 \) values of the grass and soil classes are all well above 0.70, which can be seen as a satisfactory result due to the complex nature of this task.

V. DISCUSSION

As stated earlier, the primary objective of this paper is to illustrate the advantages of examining the physical nature of very high spatial resolution image time series, including the use of surface reflectance and angular decompositions. Part of the motivation is the tendency of recent studies to account only for image DNs, ignoring the potential improvements possible using a physical approach. Machine learning studies have improved the development of complex Earth observation systems, providing valuable contributions to the advancement of remote sensing. However, these systems often overlook the benefits of understanding the physical nature of such data, with the end result of failing to take full advantage of the imagery available [18]. This is even more relevant when large image time series need to be analyzed. For instance, this study showed that the image values are affected by the atmospheric and geometric properties of the acquisitions. Therefore, the domain adaptation concept reviewed in Section I may be studied to minimize local and underdetermined effects, such as the spectral variations due to the BRDF (and the implications of working with sensors on agile satellite platforms), instead of focusing on the image-scaled distortions.

The challenge of the future is to have robust systems that can easily analyze a large number of very high spatial resolution images, making the best use of the data collected over the years. In contrast, newer techniques in the literature are usually developed on a very limited data set over a small number of classes. Part of the problem is the cost of the imagery that can still be prohibitive, representing a major issue that prevents long-term studies. The authors also realize that the community is lacking common data sets that can be used to compare the various results. On the other hand, there are a few initiatives aimed at providing high-quality data, such as the one promoted by the IEEE-Geoscience and Remote Sensing Society (GRSS) Standardized Algorithm Development and Evaluation working group, which is focused on standardizing data sets and performance measures for algorithm evaluation and comparison [56], or the one endorsed by the IEEE-GRSS Data Fusion Technical Committee, which encourages cutting-edge research of remote sensing image analysis [57]. These initiatives provide an unbiased basis for the community to evaluate and compare algorithms and results and to advance remote sensing research with a common performance metric.

VI. CONCLUSION

The availability of submeter optical imagery regularly acquired over the same geographical region has proved to be effective in a large number of applications, from precision agriculture to disaster management and to urban planning. Despite the progress in space technology in operating more sophisticated sensors and the large amount of data made available, very little research addresses the advantages and challenges of multitemporal and multiangular optical very high spatial resolution spaceborne imagery.

This paper illustrated not only that physical quantities are necessary to consistently and efficiently analyze these kinds of data sets but also that the angular information of the acquisitions should not be neglected, as unique additional features can be derived from it. More importantly, the temporal and angular components should always be simultaneously considered as some of the radiometric differences in the time series (the so-called data set shift in the machine learning terminology) may often be leveraged or accounted for by understanding
The Earth–Sun distance can be determined using the following equations [58]:

\[
d_{ES} = 1.00014 - 0.01671 \cdot \cos(g) - 0.00014 \cdot \cos(2g)
\]

with

\[
g = 357.529 + 0.98560028 \cdot A
\]

\[
A = JD - 2451545.0
\]

\[
JD = [365.25 \cdot (year + 4716)] + [30.6001 \cdot (month + 1)] + \text{day} + \frac{UT}{24 \text{ hour}} + B - 1524.5
\]

\[
UT = \text{hour} + \frac{\text{minute}}{60} + \frac{\text{second}}{3600.0}
\]

\[
B = 2 - C + \left[\frac{C}{4}\right]
\]

\[
C = \left\lfloor\frac{\text{year}}{100}\right\rfloor
\]

where \( \text{year, month, day, hour, minute, and second} \) correspond to the acquisition time of the image and \( \left\lfloor \cdot \right\rfloor \) represents the integer part of the number. Note that, if the image is acquired in January or February, then \( \text{year and month} \) must be modified as follows.

1) \( \text{year} = \text{year} + 1 \).

2) \( \text{month} = \text{month} + 12 \).

**Acknowledgment**

The authors would like to thank V. Leonard and D. Tuia for the insightful and helpful comments.

**References**


Fabio Pacifi (S’03–M’10–SM’13) received the Ph.D. degree in geoinformation and the Laurea Specialistica (M.Sc.; cum laude) and Laurea (B.Sc.; cum laude) degrees in telecommunication engineering from Tor Vergata University, Rome, Italy, in 2010, 2006, and 2003, respectively.

Since 2009, he has been with DigitalGlobe, Inc., Longmont, CO, USA, as an R&D Scientist. Between 2005 and 2009, he collaborated as a Visiting Scientist with the Department of Aerospace Engineering Sciences, University of Colorado, Boulder, CO, USA.

He has been involved in various remote sensing projects commissioned by the European Space Agency. His research activities include the processing of remote sensing images, data fusion, feature extraction, pattern recognition, and analysis of multitemporal and multitemporal data. His research interests include classification and change detection techniques of urban areas using very high spatial resolution optical and/or synthetic aperture radar imagery, with special emphasis on machine learning. He has authored (or coauthored) more than 70 scientific publications, including journal papers, book chapters, and peer-reviewed conference proceedings.

Dr. Pacifi is the current Editor of the IEEE Geoscience and Remote Sensing eNewsletter and serves as an Associate Editor for the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING (JSTARS). He served as the Chair of the IEEE Geoscience and Remote Sensing Society Data Fusion Technical Committee (2011–2013). He was the recipient of the 2011 Best Reviewer Award from the IEEE Geoscience and Remote Sensing Society for his service to IEEE JSTARS. He was also the recipient of the Best Student Paper Award at the 2009 IEEE Joint Urban Remote Sensing Event. He ranked First at the 2007, 2008, and 2009–2010 IEEE Geoscience and Remote Sensing Society Data Fusion Contest. He was the Guest Editor of a special issue of the IEEE JSTARS on multitemporal remote sensing.

**PACIFICI et al.: IMPORTANCE OF PHYSICAL QUANTITIES**


Nathan Longbotham (S’11–M’14) received the B.S. degree in physics (magna cum laude; university scholar; presidential scholar; Fred J. Barton departmental award) from Abilene Christian University, Abilene, TX, USA, in 2001, the M.S. degree in optical science and engineering from the University of New Mexico, Albuquerque, NM, USA, in 2008, and the Ph.D. degree in aerospace engineering sciences from the University of Colorado, Boulder, CO, USA, in 2012 with a remote sensing certification.

He is currently a Remote Sensing Scientist in the Research Laboratory of DigitalGlobe, a pioneer in the field of very high spatial resolution satellite imagery based in Longmont, CO, USA. His current research is specialized in remote sensing spectral science, image analysis, and data fusion of optical satellite systems. Prior to his involvement in satellite-based remote sensing, his research was in the field of laser physics and included a position with the holographic data storage start-up, InPhase Technologies, developing a specialized tunable ultraviolet laser system, as well as a graduate internship position at Sandia National Laboratories, where he conducted research into the properties of Q-switched microlasers and their applicability to Light Detection And Ranging systems.

Dr. Longbotham has been an active member of the remote sensing community. This involvement includes earning the first place in the 2009–2010 IEEE Geoscience Remote Sensing Data Fusion Contest, presenting as an invited finalist in the 2013 International Geoscience and Remote Sensing Symposium student paper competition, and frequently volunteering as a reviewer for the journal publications of the IEEE Geoscience and Remote Sensing Society.

William J. Emery (F’02) received the Ph.D. degree in physical oceanography from the University of Hawaii, Hilo, HI, USA, in 1975.

After working at Texas A&M University, College Station, TX, USA, he moved to the University of British Columbia, Vancouver, BC, USA, in 1978, where he created a Satellite Oceanography facility/education/research program. He was appointed Professor in Aerospace Engineering Sciences at the University of Colorado, Boulder, CO, USA, in 1987. He is an Adjunct Professor of Informatics at Tor Vergata University in Rome, Italy. He has authored over 175 refereed publications and two textbooks.

Dr. Emery is a Fellow of the IEEE and the Vice President for publications of the IEEE Geoscience and Remote Sensing Society (GRSS). He was the recipient of the GRSS Educational Award in 2004 and its Outstanding Service Award in 2009. He is a Fellow of the American Meteorological Society (2010) and the American Astronautical Society (2011) and was recently elected a Fellow of the American Geophysical Union (2012).