Spectral variability of remotely-sensed target materials:
Causes, models, and strategies for mitigation and robust exploitation

J. Theiler, A. Ziemann, S. Matteoli, and M. Diani

Abstract—The central problem in hyperspectral remote sensing is characterizing the material components of a scene based on the spectral radiance observed in the image pixels. What makes this challenging is that the spectral response for a given material exhibits considerable variability from a variety of causes: intrinsic (depending on composition or morphology of the material), extrinsic (depending on size of an object, or concentration of the material), or environmental (due to illumination, atmospheric distortion, *etc.*). In this article, we survey many of the causes of spectral variability, describe spectral models for this variability, and outline some signal processing and target detection strategies for analyzing hyperspectral data in a way that is more robust to this variability.

I. HYPERSPECTRAL IMAGING

Many remote sensing applications require locating a specific target object or material within a scene, which may then be further identified, characterized, or quantified. A geologist might care about a particular mineral; a forester might want to identify and monitor certain vegetation species; a public health agent might want to know about particular pollutants and dangerous toxins in the environment; a climate scientist might want to locate sources of methane or carbon dioxide; or a government might want information about the nuclear proliferation activities in another country.

With exquisite discrimination ability enabled by upwards of hundreds of distinct spectral channels in every pixel, hyperspectral imaging has the potential to solve all of these problems. Underlying the many algorithms developed for exploiting hyperspectral imagery is the notion that each material is unequivocally characterized by its unique spectral signature (embodied by the spectral reflectance, in the visible, near infrared, and shortwave infrared; or by the spectral emittance, in the longwave infrared).

Nonetheless, the notion of a single predetermined spectral signature for each material is an ideal concept that is not observed in real-world applications. Even

J. Theiler and A. Ziemann are with Los Alamos National Laboratory, Los Alamos, NM 87545 USA (e-mail: {jt,ziemann}@lanl.gov). S. Matteoli is with the National Research Council of Italy (CNR) – Institute of Electronics, Computer and Telecommunication Engineering (IEIIT), Italy (e-mail: stefania.matteoli@ieiit.cnr.it). M. Diani is with the Italian Naval Academy, Livorno, Italy (e-mail: m.diani@iet.unipi.it).

in precisely-controlled laboratory experiments, the measured spectral reflectance of the same material exhibits some variability. Matters get even worse in hyperspectral remote sensing where target materials are imaged by a sensor mounted on a platform, as many environmental and geometric aspects are involved and unavoidably affect the measured spectral radiance for the given material. All these sources of variability will alter the spectral intensity and shape of a measured spectrum and will affect the performance of hyperspectral image exploitation, particularly in the case of target detection.

The purpose of this survey is to examine the major causes of spectral variability in target materials, to survey some of the associated physical models, and to discuss how hyperspectral analyses can be made more robust to this variability. Whereas there are many kinds of hyperspectral image exploitation algorithms (*e.g.*, classification, clustering, unmixing) [1], and virtually all of them are affected by spectral variability, we will concentrate on the problem of target detection and, more specifically, the detection, identification, and location of particular materials of interest.

> *"I don't like sand. It's coarse and rough and irritating and it gets everywhere."*
> — Anakin Skywalker

II. SOURCES OF TARGET SPECTRAL VARIABILITY

The most obvious and immediate cause of variability in a target spectrum is the intrinsic variability of the target material itself. Unless the material is a pure substance, there is likely some variability in its composition; for instance, we can refer to "sand" as a single material, but there are many kinds of sand, and any given kind of sand is composed of varying ratios of mineral components [2]. Even "pure" materials can undergo chemical variation (*e.g.*, due to oxidation or hydration), and even a chemically pure material, with fixed optical properties, can exhibit spectral reflectances that vary with material morphology. In addition to these intrinsic variability sources, the effective reflectance at a pixel can depend on extrinsic properties of the material, such as concentration or thickness or size relative to the pixel. Environmental effects, although not affecting the target reflectance itself,
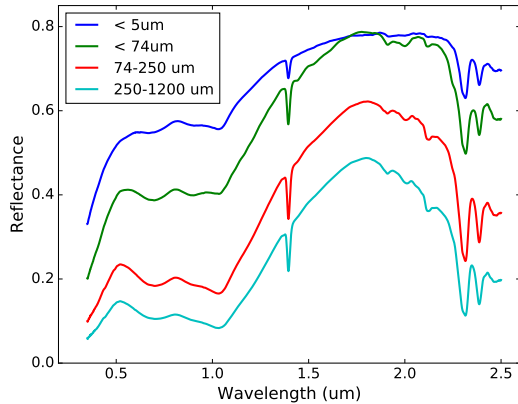
Fig. 1. Four spectra of the mineral actinolite, based on four ranges of particle sizes, with the smaller particles generally more reflective. From the USGS database [3].

Fig. 2. Illustration of BRDF geometry described in Eq. (1).

can have a large impact on observed at-sensor spectral radiance; these effects include atmosphere, illumination, adjacent materials, and issues within the sensor itself.

The ultimate consequence of all these sources of variability is that the observed target spectrum in the field often differs from spectral reflectance or emissivity measured in the laboratory and archived in spectral databases, such as the United States Geological Survey (USGS) Spectral Library [3]. The first step in bridging this gap is to identify and catalog these individual sources, while at the same time recognizing that their effects occur together and superimpose upon each other. If the sources are understood, then models can be formulated, and from these models, algorithms can be designed to achieve some measure of robustness to this variability.

A. Intrinsic target variability

The most direct source of spectral variability in a target material is due to the chemistry or morphology inherent to the material. When in powdered form, "pure" solid materials become particularly complicated as their spectra are highly variable [4] depending on particle size (*e.g.*, see Fig. 1), particle shape, and packing density, and can be further complicated in the case of fine-grained (*a.k.a.* "intimate") mixtures of multiple materials. Under ideal conditions, however, the fundamental reflectance properties of a pure solid material are directly characterized by its optical constants (refraction index n and extinction coefficient k) [5], [6]; from those values one can derive modeled reflectance signatures based on morphology (*e.g.*, for particle size [7], [8] or packing density [9], [10]). As sensor technology has improved in recent years, there has been an increasing need for such reflectance models in order to generate *simulated* spectra, as obtaining comprehensive spectral measurements under
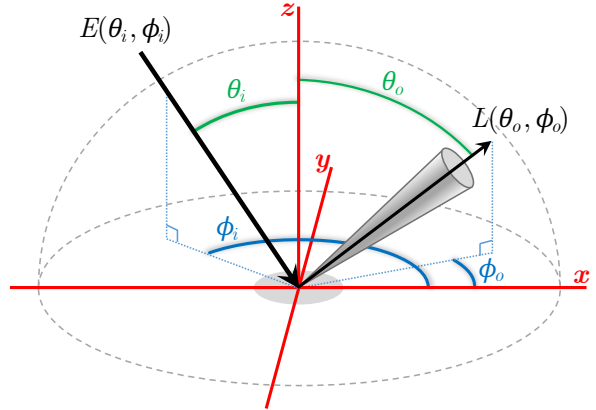
the multitude of possible morphologies is intractable [11], [12]. For solid materials, both pure and mixed, this reflectance also depends on both the angle from which it is illuminated and the angle from which it is observed. Those dependencies are encapsulated in the Bidirectional Reflectance Distribution Function (BRDF) [13]–[15], which is a function of four variables that defines how light reflects off of an opaque surface [16]:

$$r_{BRDF} = \frac{L(\theta_o, \phi_o)}{E(\theta_i, \phi_i)} \quad [sr^{-1}]. \tag{1}$$

Here, the bidirectional reflectance r_{BRDF} is the ratio of the radiance $L(\theta_o, \phi_o)$ to the irradiance $E(\theta_i, \phi_i)$, where L is scattered into the direction described by the orientation angles θ_o and ϕ_o, and E is irradiance from the θ_i, ϕ_i direction (see Fig. 2 for an illustration of BRDF geometry). The function in Eq. (1) describes the bidirectional reflectance values for every combination of input/output angles, and, although not explicitly captured in the equation here, also changes as a function of wavelength.

The simplest BRDF is constant with respect to the input and output angles; such a surface is termed *Lambertian*.

Although the assumption of an ideal diffusely-reflecting surface may not truly be achieved in practice, it is often a useful approximation, and a natural starting point for more sophisticated models. Some of these BRDF models are theoretical physics-based models based on simulations of light scattering from multi-faceted and/or multi-layered reflecting rough surfaces, while others are empirical models that provide simple formulations capable of reproducing specific kinds of reflective behaviors. Often these models involve free parameters that are fit to comprehensive experimental measurements made using sophisticated goniometers and gonioreflectometer systems; the challenge there is that

those systems are only recently being designed for practical deployment in challenging environments [17], [18]. The more complex models are often employed in computer graphic design to render realistic surface illustrations [19], while simpler models are more commonly adopted in remote sensing. Schowengerdt [20] argues that most natural surfaces are approximately Lambertian for viewing angles (*i.e.*, θ_o in Fig. 2) within 20° to 40° from the zenith. But BRDF also depends on the illumination angle, with most surfaces brighter in the specular and backscatter directions and darker elsewhere. Trees and canopies, for instance, have a strong return from the backscatter direction [16]. Choosing the Lambertian model certainly simplifies hyperspectral data analysis, but it just as clearly under-estimates the spectral variability that will be observed in both target and background materials.

Intimate mixtures present their own set of challenges, as their spectra result from complicated and varying underlying phenomenology (*e.g.*, reflections and scatterings between grains of different materials). A common model for these nonlinear spectral interactions is Hapke's radiative transfer model for intimate mineral mixtures [21], [22]. This model considers several contributions to the spectra of intimate mixtures: single volume scattering, multiple volume scattering, coherent backscattering, the shadow hiding opposition effect, macroscopic roughness, and compactness. All of these contributions can result in highly variable resulting spectra. A number of approaches have been explored for applications to hyperspectral unmixing, with particular focus on kernel-based methods [23]–[27].

B. Extrinsic target variability

The signal observed at the sensor depends not only on the nature of the material, but also on *how much* material there is. For solid sub-pixel targets, the target size relative to the pixel size matters; for gas-phase plumes or thin layers of powder, concentration or thickness matters. In the long-wave infrared (LWIR) the temperature matters as well (warmer materials emit more radiation, and appear brighter at the sensor).

Perhaps the simplest, and probably the most common, expression for how much material is in a pixel is given by the *additive* model [1]:

$$\mathbf{x} = \mathbf{z} + \epsilon \mathbf{t} \tag{2}$$

where $\mathbf{x}$ is the observed spectrum at a pixel, $\mathbf{z}$ is the spectrum of the background material in that pixel, $\mathbf{t}$ is the spectral signature of the target material, and ϵ is a scalar quantity that corresponds to how much of the target is present.

For opaque subpixel targets, the larger the target, the more of the background it obscures, and a more appropriate expression is the *replacement* model [1]:

$$\mathbf{x} = (1 - \alpha)\mathbf{z} + \alpha \mathbf{t} \tag{3}$$

where $0 \leq \alpha \leq 1$ corresponds to the fraction of the pixel that is taken up by the target material.

While the additive and replacement models are the most popular choices, the actual interaction of target and background can be more complicated. For chemical plumes, for instance, the additive model is popular, but it is only an approximation – accurate in the weak-plume limit [28] – to the actual Beer's law absorption rule [29]. This approximation is often valid, but it is also motivated by the fact that target detection algorithms are more easily derived under the assumptions of the additive model.

The *scaled replacement* model, which is a kind of hybrid of additive and replacement, considers a target that can both occlude the background and have variable strength:

$$\mathbf{x} = (1 - \alpha)\mathbf{z} + \alpha \epsilon \mathbf{t}. \tag{4}$$

In the LWIR, this could for instance be due to temperature variation. An example in the visible wavelengths arises when a target in the shade exhibits a lower *apparent* reflectance than that same target in the bright sun. More sophisticated models (*e.g.*, that reduce the dependence on the temperature [30] or achieve temperature-emissivity separation [31], or that explicitly account for shadows [32]) may ultimately be preferable to the scaled replacement, but they are also more complicated.

C. Environmentally-induced target variability

Target spectral variability does not only depend on the target itself but is also caused by the surrounding environment. In this subsection, we examine some of the most important environmental sources of spectral variability by considering the typical remote sensing scenario of a hyperspectral sensor that is mounted on board a platform (such as an aircraft or a satellite) and is imaging the targeted surface materials.

1) Main radiative transfer model equations: The radiative transfer between a surface target and a sensor across the visible and near infrared (VNIR, $\sim$400-1000nm) and short-wave infrared (SWIR, $\sim$1000-2500nm) spectral ranges (*i.e.*, VNIR-SWIR) can be expressed by the following simplified physics-based model for the sensor-reaching radiance $L^s(\lambda)$ [16], [33]–[36]:

$$L^s(\lambda) = L^{gr}(\lambda) + L^{sc}(\lambda) \tag{5}$$

$$L^{gr}(\lambda) = L^{dir}(\lambda) + L^{dif}(\lambda) + L^{obst}(\lambda) \tag{6}$$

$$L^{sc}(\lambda) = L^p(\lambda) + L^{adj}(\lambda). \tag{7}$$

A sketch of the various radiation contributions entering the sensor is provided in Fig. 3, with Table I summarizing the major radiometric terms involved. In the above equations, the total ground-reflected spectral radiance $L^{gr}(\lambda)$ includes the contributions that are reflected by the target towards the sensor, and the solar scattered spectral radiance $L^{sc}(\lambda)$ includes the contributions that are scattered towards the sensor without having interacted with the targeted surface. $L^{dir}(\lambda)$ and $L^{dif}(\lambda)$ are the major contributors to $L^{gr}(\lambda)$ and represent, respectively, the direct and diffuse (downwelling) solar spectral radiances reflected by the target towards the sensor – also accounting for the multiple reflections between the adjacent material surface and the atmosphere due to multiple scattering [16]. The ground reflected radiance also includes the spectral radiance $L^{obst}(\lambda)$ arriving on the target due to *secondary illumination* from potential nearby objects and obstacles in view of the target (such as trees and buildings) [35]. The solar scattered radiance $L^{sc}(\lambda)$ includes the so-called *path radiance* $L^{p}(\lambda)$, *i.e.*, the solar spectral radiance that is scattered by the atmosphere towards the sensor, and the adjacency term $L^{adj}(\lambda)$, accounting for the direct and diffuse solar spectral radiances reflected from the materials adjacent to the target towards the sensor (including the aforementioned multiple reflections phenomenon).

By expanding the first two terms on the right hand side of Eq. (6), we obtain [16], [33], [36]:

$$L^{dir}(\lambda) + L^{dif}(\lambda) = \frac{E^{su}(\lambda) + F \cdot E^d(\lambda)}{1 - r_a(\lambda)S(\lambda)} \frac{r(\lambda)}{\pi} T^u(\lambda) - \alpha_{sh} E^{su}(\lambda) \frac{r(\lambda)}{\pi} T^u(\lambda), \tag{8}$$

$$E^{su}(\lambda) = E^{toa}(\lambda) \cos\theta \, T^d(\lambda). \tag{9}$$

In Eqs. (8,9), $E^{su}(\lambda)$ is the solar unscattered (direct) irradiance incident on the surface, with $E^{toa}(\lambda)$ as the solar exo-atmospheric spectral irradiance at the top of atmosphere incident on a surface orthogonal to the Sun's rays, θ as the Sun zenith angle (subtended by the Sun's rays and the surface normal), and $T^d(\lambda)$ as the downward atmospheric transmittance of the Sun-to-surface path. The term $E^d(\lambda)$ is the downwelling spectral irradiance from the sky incident on the surface, which undergoes a scaling by the *sky view factor* (or *shape factor*) $F \in [0, 1]$ accounting for the fraction of sky-dome visible from the target [35], [36]; $r(\lambda)$ is the diffuse spectral reflectance (or *albedo*) of the targeted surface (assumed Lambertian; *i.e.*, $r(\lambda) = r_{BRDF}(\lambda) \cdot \pi$); $r_a(\lambda)$ is the diffuse spectral reflectance of the material adjacent to the target; $S(\lambda)$ is the atmospheric spherical scattering albedo (*i.e.*, the effective diffuse reflectivity of the sky to upwelling radiation); and $T^u(\lambda)$ is the upward atmospheric transmittance of the target-to-sensor path. The denominator in the first term of Eq. (8) is due to the multiple reflections between the adjacent material surface and the atmosphere from multiple scattering. This phenomenon, also known as the "trapping effect," has been modeled for a flat and homogeneous surface as the summation of infinite terms of reflected contributions leading to a geometric series that converges to $[1 - r_a(\lambda)S(\lambda)]^{-1}$. This effect has been found negligible for clear sky conditions and low reflectivity of adjacent materials (*i.e.*, $r_a(\lambda)S(\lambda) < 0.02$ [16]). The second term in Eq. (8) accounts for potential shadowing of the direct solar unscattered term, with $\alpha_{sh} = 1$ for full shadowing and $\alpha_{sh} = 0$ for full illumination [33].

From the point of view of spectral remote sensing, the factor we care most about in these equations is $r(\lambda)$, the reflectance. This is what corresponds to the intrinsic properties of the material on the ground, *i.e.*, the material we are trying to find or identify or characterize. Extracting $r(\lambda)$ from the sensor-reaching radiance $L^s(\lambda)$ is no small feat, as in doing so we not only have to estimate all of the other terms and factors in Eqs. (5-9), but we also have to deal with the variability of each of those terms and factors.

The terms $L^{obst}(\lambda)$ and $L^{adj}(\lambda)$ in Eqs. (6,7) are not expanded here, but details can be found in Refs. [16], [35], [37], [38]. $L^{adj}(\lambda)$ has basically the same form as Eq. (8), with $r_a(\lambda)$ in place of $r(\lambda)$ and with the diffuse upward transmittance instead of $T^u(\lambda)$. As to $L^{obst}(\lambda)$, this has been modeled taking into account that the presence of a nearby object that obstructs part of the sky-dome viewable from the target and, thus, the average spectral radiance reflected by the object itself is scaled by a factor $(1 - F)$ [16], [35]. Depending on the relative position of the Sun, target, and obstacle, $L^{obst}(\lambda)$ may also include the solar direct radiation reflected by the nearby object towards the target.

a) Topographic effects: So far we have assumed a flat surface, but topographic effects and their possible incorporation into the model have been thoroughly studied (see [34], [37], [39]–[42]), and Eqs. (8,9) can be modified accordingly. Specifically, for a local target surface at location (x, y) tilted by a slope angle $\theta_{sl}(x, y)$ (subtended by the surface normal and the vertical direction), Eqs. (8,9) should be modified by accounting for the local direct and diffuse irradiances at (x, y):

$$E^{su}(\lambda; x, y) = E^{toa}(\lambda) \cos[\theta_{ill}(x, y)] T^d(\lambda) \tag{10}$$

$$F(x, y) \cdot E^d(\lambda; x, y) = E^d(\lambda) \cdot \left\{ A(\lambda) \frac{\cos[\theta_{ill}(x, y)]}{\cos\theta} + [1 - A(\lambda)] F(x, y) \right\} \tag{11}$$
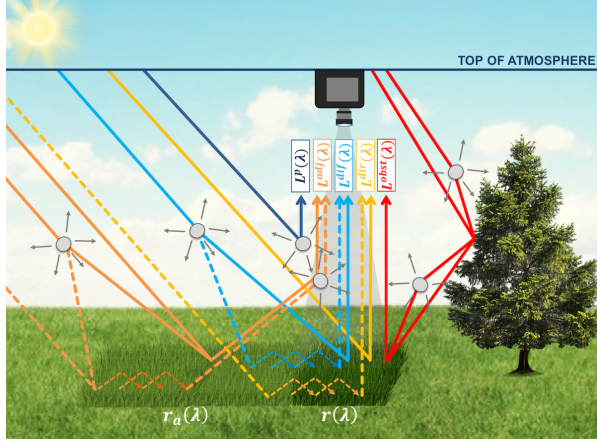
Fig. 3. Sketch of the radiative transfer model along the Sun-to-target-to-sensor path, as described in Section II-C1. The individual terms are listed in Table I.

Spectral Radiance	
$L^s(\lambda)$	Sensor reaching spectral radiance
$L^{gr}(\lambda)$	Total ground reflected spectral radiance
$L^{sc}(\lambda)$	Solar scattered spectral radiance
$L^{dir}(\lambda)$	Direct solar spectral radiance
$L^{dif}(\lambda)$	Diffuse (*downwelling*) solar spectral radiance
$L^{obst}(\lambda)$	Nearby obstacles/objects spectral radiance (*secondary illumination*)
$L^p(\lambda)$	Solar path scattered radiance (*path radiance*)
$L^{adj}(\lambda)$	Adjacency spectral radiance
Spectral Irradiance	
$E^{toa}(\lambda)$	Exo-atmospheric spectral irradiance and the top of atmosphere
$E^{su}(\lambda)$	Solar unscattered (*direct*) spectral irradiance incident on the surface
$E^d(\lambda)$	Downwelling spectral irradiance from the sky incident on the surface
$E^{su}(\lambda; x, y)$	Solar unscattered spectral irradiance incident on a tilted surface @ (x, y)
$E^d(\lambda; x, y)$	Downwelling spectral irradiance incident on a tilted surface @ (x, y)
Spectral transmittance	
$T^u(\lambda)$	Upward atmospheric transmittance (target-to-sensor path)
$T^d(\lambda)$	Downward atmospheric transmittance (Sun-to-surface path)
Spectral reflectance	
$r(\lambda)$	Spectral reflectance of the target (assumed *Lambertian*)
$r_a(\lambda)$	Spectral reflectance of the adjacent material
$S(\lambda)$	Atmospheric spherical scattering albedo

Here $\phi_{sl}(x, y)$ is the azimuth angle of the local sloped terrain at (x, y), ϕ is the Sun azimuth angle, and $\theta_{ill}(x, y)$ is the local (actual) target illumination angle. In Eq. (11), the diffuse component is expressed as a linear combination of two terms: one circumsolar diffuse irradiance from the solid angle near the Sun and one isotropic contribution from the remaining sky-dome [37], [39], [40]. The term $A(\lambda)$ is the anisotropy index giving the proportion of diffuse radiation to be treated as circumsolar and isotropic and has been modeled as $A(\lambda) = (1 - \alpha_{sh})T^d(\lambda)$ [37], [39], [40]. The sky-view factor for the tilted target surface in Eq. (11) should be evaluated from digital elevation models (DEMs) [40], [43], though simplified expressions have also been derived (*e.g.*, for an infinitely-long slope [40]). It should be noted that Eqs. (10-12) can be modified or enriched in order to account for topography at different levels such as for mountainous terrain [37] or for the combined effect of topography and structured forests [41], [42]. Additionally, some works employ a sky radiance fraction function instead of scaling the downwelling irradiance by the sky-view factor in Eqs. (8,11), in turn obtaining a wavelength-dependent and angle-dependent term accounting for the non-Lambertian nature of the sky reflectance [44].

As shown in the equations above, many factors are involved in the radiative transfer process. Each of these factors affects the spectral intensity and shape of the sensor-measured radiance spectrum, and can ultimately result in a significant amount of spectral variability. The sources of variability may be broadly categorized into four areas: atmosphere, illumination, acquisition geometry, and adjacent environment. In the following section, a detailed analysis of the effects of this variability are examined based on the radiative transfer model equations just described.

2) Analysis of environmentally-induced variability: Most of the variability in an observed spectrum is due to atmosphere and illumination, and these sources of

variability have indeed been the most widely investigated [36], [45]–[48]. Many have shown, either by simulation or by field experiments, how much the spectral radiance for a single material can vary under different atmospheric conditions (*e.g.*, aerosol/gas types and concentrations, or water vapor profiles) [46] and illumination conditions (*e.g.*, shadowing, or sky-dome obstruction) [36], [45], [47], [48]. In order to investigate adjacency effects and secondary illumination, Goa *et al.* [48] looked at radiance measurements of targets that were placed between tree lines and exposed to different illumination conditions; these measurements were then compared to synthetic radiance spectra simulated according to a model based on Eqs. (5-7). Results revealed the importance of accounting for the secondary illumination term $L^{obst}(\lambda)$, which is often neglected, and showed that it can play an important role in cases where, for instance, the target is in shadow and surrounded by tall and brightly lit objects [35]. Similar considerations can be made for the effects of adjacent materials both in terms of $L^{adj}(\lambda)$ and $[1 - r_a(\lambda)S(\lambda)]$. They may not be negligible, especially when $r_a(\lambda) \gg r(\lambda)$ [37], *i.e.*, when adjacent materials are more reflective than target materials (such as targets surrounded by brighter backgrounds and, in particular, targets in shadow) [35], [49]. Viewing conditions and acquisition geometry also play an important role in spectral variability of materials [50], [46], especially topographic effects that may considerably increase or decrease the direct and diffuse illumination of the target. For example, consider a Sun zenith angle of 25°, a slope angle of the target surface of 15° away from the Sun, and a null azimuthal difference between the Sun and the slope orientation. In this case, the direct illumination would decrease about 15% compared to the flat terrain case. For the same geometry but with a Sun zenith angle of 45°, the decrease would be about 30%.

The atmospheric effects of transmittance and visibility in general tend to vary slowly with respect to spatial position, and are often assumed to be stationary over a full scene. This assumption can fail for wide-area remote sensing or with shallow slant-angled views (*oblique sensing*) [50], [51], where the target-to-sensor path distance may vary considerably over a scene. Other sources of variability, such as columnar water vapor [52], clouds, shadows, and obstacles (*e.g.*, leafy canopies and buildings) can be quite spatially non-stationary. Topographic effects also exhibit a strong spatial non-stationarity within a scene. In part, this is because of the increased range of sensor viewing angle within a scene [50], [51], but a further effect of oblique sensing is that surfaces that were previously hidden become visible (*e.g.*, walls of a building as well as the roof, or trunks of a tree as well as the canopy) [51]. Adjacency effects are intrinsically non-stationary, of course, as they depend on

the spatial variations of surface materials in the examined scene.

A summary of the considerations above is reported in Table II. Graphical examples showing how these sources of variability manifest themselves on the sensor-reaching radiance are provided in the following section.

a) Illustrative examples of environmentally-induced variability effects: An example experiment is shown in Fig. 4, where at-sensor radiance spectra in the VNIR were synthetically generated by exploiting spectral signatures from the ASTER spectral library [53]. The spectral reflectance $r_{Cu}(\lambda)$ of *copper* was used for the target material. This was chosen because it has a smooth reflectance spectrum and exhibits both low reflectivity (in the blue/green portions of the VIS) and medium/high reflectivity (in the red and NIR). To illustrate the effects of adjacent and obstacle materials with respect to the copper, we used a *rangeland* reflectance spectrum $r_{Rl}(\lambda)$ with very low reflectivity across the full VNIR; an *olive canopy* reflectance spectrum $r_{Ol}(\lambda)$ with high reflectivity in the NIR only; and an *aluminum* reflectance spectrum $r_{Al}(\lambda)$ with higher reflectivity across the full VNIR. A plot of these four spectra is shown in Fig. 4(a).

In this example, the MODerate resolution TRANsmission (MODTRAN) 5 radiative transfer code [54] was employed for simulating radiation transfer in the atmosphere in combination with the physical models expressed by Eqs. (10-12). A typical airborne remote sensing scenario was reproduced by making reference to a real hyperspectral data collection campaign performed in May 2013 with the SIM.GA hyperspectral sensor (511 spectral channels in the VNIR with a full width at half maximum [FWHM] of about 2 nm) flying over the city of Viareggio, Italy [55]. The employed parameters are listed in Table III, where the values of the varied parameters are reported in curly brackets. One single parameter was varied each time by keeping the others fixed at the basic configuration value highlighted in bold red. In Fig. 4(b-h), the spectra corresponding to the basic configurations are displayed in blue in each plot. The other colors are assigned to the variations expressed in each plot legend.

Specifically, we varied: the horizontal visibility (the surface meteorological range VIS parameter in MOD-TRAN) in Fig. 4(b); the water vapor (WV) scaling factor (H2OSTR in MODTRAN) in Fig. 4(c); the sensor viewing angle θ_v (ANGLE in MODTRAN) in Fig. 4(d); shadowing (by means of α_{sh}) in Fig. 4(e); topography (by taking $\phi - \phi_{sl} = 0$ and varying θ_{sl}) in Fig. 4(f); adjacency effects (using a null albedo or olive or rangeland or aluminum signature for the adjacent material) in Fig. 4(g); and obstacles and obstruction (employing a null albedo or olive or aluminum signature for both 15% and 30% sky-dome obstructions) in Fig. 4(h). With

Category	Sources of variability	Mainly affected terms	Spatial stationarity	Temporal stationarity	Category
	Constituents, gases, aerosols	$E^d(\lambda), L^p(\lambda), T^d(\lambda),$ $T^u(\lambda), S(\lambda)$	Stationary over wide areas*	Seasonality	ATMOSPHERE
	Clouds, water vapor	$T^d(\lambda), T^u(\lambda), L^p(\lambda),$ $E^d(\lambda), S(\lambda)$	Non-stationary	Non-stationary	ATMOSPHERE
ADJACENT ENVIRONMENT	Adjacent materials	$L^{adj}(\lambda), r_a(\lambda)$	Non-stationary	Stationary	
ADJACENT ENVIRONMENT	Nearby objects and obstacles	$L^{dif}(\lambda), L^{obst}(\lambda)$	Non-stationary	Stationary	ILLUMINATION
ADJACENT ENVIRONMENT	Shadows	$L^{dir}(\lambda)$	Non-stationary	Hourly variations	ILLUMINATION
ACQUISITION GEOMETRY	Sun position	$E^{su}(\lambda),$ $E^d(\lambda), L^{obst}(\lambda),$ $L^{adj}(\lambda) \ T^d(\lambda), T^u(\lambda)$	Stationary over wide areas	Hourly variations	ILLUMINATION
ACQUISITION GEOMETRY	Topography	$E^{su}(\lambda), L^{dif}(\lambda),$ $L^{obst}(\lambda), L^{adj}(\lambda)$	Non-stationary	Stationary	ILLUMINATION
ACQUISITION GEOMETRY	Sensor viewing angle	$T^u(\lambda), L^p(\lambda)$	Stationary**	Stationary	

*except for wide imaged areas and situations with extremely wide-angle views, or very shallow slant angled views
** except for situations with extremely wide-angle views

respect to the basic configuration of VIS = 23km (the typical default value in MODTRAN for good visibility conditions), visibility variations strongly impact the spectra, as shown in Fig. 4(b). The major effects can be observed in the low wavelength region for low VIS values, where the low target reflectivity combined with the increased scattering phenomena due to the hazier conditions results in considerable increases of the at-sensor spectral radiance. For higher wavelengths, where not only is scattering generally weaker but the copper target also has higher reflectance, the general effect on the at-sensor radiance is a decrease of the signal for low VIS values due to the less transparent atmosphere (lower transmittance). The result is a completely distorted signal. In contrast, in Fig. 4(c) the effects of WV can be seen to be mostly confined to the WV absorption bands, with decreased signal (more absorption) for more humid atmospheres (higher WV scaling). Oblique sensing (Fig. 4(d)) plays a significant role only for very slanted views (e.g., $\theta_v = 120°$), which mostly impact lower wavelengths with an increased at-sensor radiance due to the increased scattering phenomena occurring in the longer target-to-sensor atmospheric path. Very oblique views also impact, though to a lesser extent,

the signal in the major absorption bands (e.g., at 940 nm) where decreased signal intensities can be observed due to the higher absorption. The reduction of the direct solar illumination due to shadowing entails considerable variability over the at-sensor spectral radiance in Fig. 4(e), with a general decrease of the at-sensor signal throughout the VNIR range. Similar variations can be observed in Fig. 4(f), where varying the inclination of the targeted surface determines a general increase (if the surface is rotated towards the Sun) or decrease (if it is rotated away from the Sun) of the at-sensor signal. In the simulated scenario the Sun zenith angle is $\theta = \theta_i \approx 31°$ and thus an inclination of $\theta_{sl} = 30°$ is close to the condition of maximum illumination due to topography.

The adjacency effects shown in Fig. 4(g) illustrate how much those effects depend on the properties of the adjacent material itself. In this example, the low-reflectivity rangeland spectrum hardly impacts the resulting at-sensor radiance. Conversely, for adjacent materials with stronger reflection properties, such as the olive or the aluminum materials, adjacent effects are more significant. The olive signature "colors" the at-sensor radiance according to its own spectral characteristics, with adjacency effects mostly manifesting themselves

TABLE III
PARAMETERS EMPLOYED IN THE EXAMPLE IN FIGURE 4

ATMOSPHERIC MODEL	MidLatitude Summer	VISIBILITY [Km]	$\{5, 10, \mathbf{23}, 35\}$
AEROSOL MODEL	Maritime	WV SCALING FACTOR (WV_{sf})	$\{0.5, \mathbf{1}, 2\}$
SENSOR ELEVATION [Km]	1.250	ADJACENCY	$r_a = \{\mathbf{0}, r_{Ol}, r_{Rl}, r_{Al}\}$
TARGET ALTITUDE [Km]	0	OBSTACLES and OBSTRUCTION	$r_{obst} = \{\mathbf{0}, r_{Ol}, r_{Al}\}$
			$F = \{\mathbf{0}, 0.85, 0.70\}$
TARGET LATITUDE [°]	43.85 North	VIEWING ANGLE (θ_v) [°]	$\{120, 140, 160, \mathbf{180}\}$
TARGET LONGITUDE [°]	10.25 East	SHADOWING (α_{sh})	$\{0, 0.5, \mathbf{1}\}$
DAY	May 9 2013	SLOPE ELEVATION (θ_{sl}) [°]	$\{\mathbf{0}, \pm 15, \pm 30\}$
GMT TIME [h]	12.65	SL. REL. AZIMUTH ($\phi_{sl} - \phi$) [°]	0

in the NIR spectral range; in the VIS these effects are negligible, as the olive material and copper have a null contrast in that range of the spectrum. The aluminum object, which has higher reflectivity than copper throughout the spectral range, causes a general increase in the at-sensor signal. Similar considerations can be found for the secondary illumination due to nearby objects (see Fig. 4(h)), which cannot be considered negligible for high-reflectivity objects. Here the shape factor induces variability as well, with an increasing secondary illumination impact for higher obstruction levels.

Although BRDF effects are intrinsic to the target material (and have been described in section II-A), spectral variability due to BRDF also has an environmental flavor. Indeed, variability due to variations in illumination and acquisition geometry is deeply intertwined with variability due to the non-Lambertian nature of target materials. Thus, in order to provide some insight into environmentally-induced spectral variability for non-Lambertian targets, a further illustrative example, using the same parameters as previously, is shown in Fig. 5. Two glossy targets taken from the Cornell BRDF data base [56] are considered here: specifically, a target painted with a 'Garnet red' lacquer (top panels, Fig. 5(a,b)) and a glossier target painted with a 'Mystique' lacquer (bottom panels, Fig. 5(c,d)). Here, the left panels, in Fig. 5(a,c), plot the target BRDFs for a fixed illumination direction (as before, $\theta_i = \theta \approx 31°$) for varying orientation angle θ_o. Here, the values of θ_o were chosen to correspond to the same sensor-viewing angles $\theta_v = 180° - \theta_o$ employed in the previous example. Note that the oblique specular geometry ($\theta_v = 149°$) is not included in the analysis, since the Cornell BRDF data do not include measurements of specular reflectance (if needed, however, they could be introduced by adding a term ruled by the Fresnel reflectance [57]). As is evident

from the target BRDFs, the non-Lambertian effects are different for the different targets. In Fig. 5(a), the effect is mostly just a scaling of the reflectance function for the 'Garnet red' target, with higher reflectance for nadir viewing. But in Fig. 5(c), the effect produces actual changes of color for the 'Mystique' target, which exhibits a strong reflectance contribution around 500 nm for nadir and 20° off-nadir views and strong contributions at shorter and higher wavelengths for the 40° and 60° off-nadir views. Fig. 5(b,d) plot the sensor-reaching radiance for the sensor-viewing angles matching the orientation angles varied in Fig. 5(a,c). The figures clearly show that environmental effects blend with the BRDF effects in the at-sensor spectral radiances. In fact, whereas the mostly path-radiance-driven effects already observed in Fig. 4(d) (*i.e.*, increased radiance especially at shorter wavelengths for longer paths due to oblique sensing) are clearly evident (and stronger with respect to the previous examples due to the very low reflectance of the targets), these are combined with the effects of the different spectral characteristics that the target materials intrinsically have for different viewing angles. By looking at Fig. 5(b), the path-radiance-driven spectral radiance increase at shorter wavelengths, strongly evident for the most oblique case ($\theta_v = 120°$), is less evident for the other off-nadir viewing angles because the higher overall reflectance for a nadir view (evident in Fig. 5(a)) determines a stronger radiance contribution even though the increase due to path radiance is lower than the off-nadir cases. Similar considerations can be made for Fig. 5(d), where the stronger reflectance contributions around 500 nm for the 'Mystique' target at nadir and 20° off-nadir views provide to radiance contributions around 500 nm that are stronger than or equal to the 40° off-nadir case.

These simple examples demonstrate that the causes of non-negligible spectral variability are multi-fold, and

the corresponding effects on the target spectra manifest themselves in complicated ways that can go beyond straightforward scaling or offsets. The aggregation of multiple diverse environmental effects – in addition to the spatial non-stationarity and temporal non-stationarity of the sources of these effects – makes environmentally-induced spectral variability a true challenge for effective hyperspectral image exploitation.

So far we have examined the major sources of environmentally-induced variability occurring when spectral radiance is measured in the VNIR-SWIR range. If data collection takes place with a sensor operating in the LWIR, a different radiation transfer model should be considered [58], [59] and a further source of variability comes into play, which is the temperature (primarily of the surface, but also of the atmosphere and the adjacent areas), and which exhibits further temporal and spatial non-stationarity.

We conclude this section, which is focused on the main sources of variability for the sensor-reaching spectral radiance $L^s(\lambda)$, with a reminder that $L^s(\lambda)$ as expressed here is the radiance *entering* the sensor, which is not precisely equal to the spectral radiance actually *measured* by the sensor. While beyond the scope of this survey, radiance collection by the sensor acquisition system can introduce *further* variability on the measured signal (*e.g.*, due to sensor noise, smile and keystone effects, pixel non-uniformity, point spread function, and lens distortions) [1], [60].

III. MODELS FOR SPECTRAL VARIABILITY

In attempting to capture the complex spectral variability that is observed in target materials, most models fall into two classes: statistical methods, which involve fitting a probability distribution function (pdf) to the variability; and geometrical methods, which invoke structures such as subspaces, simplexes, and manifolds. Some of these are sketched in Fig. 6, and described in a little more detail in the subsections below.

> *"Never tell me the odds."*
> — Han Solo

A. Probabilistic models

Spectral variability amounts to the same thing as spectral uncertainty, and the most natural tool for characterizing uncertainty is the probability distribution function (pdf). When *ab initio* physical models are lacking, generic distributions are appealing, and for characterizing complex backgrounds, probabilistic models are the dominant choice [61]. For characterizing the spectral variability of specific target materials, however, we often have more physical insight (*e.g.*, as detailed

in Section II). Incorporating that physical insight into algorithms is often a challenge, however, and probabilistic models are widely employed. The Gaussian pdf has many advantages [62], and several authors [63], [64] have considered a Gaussian target variability that matches (except for offset and scale) the covariance of the background variability. There may not be a strong physical argument to expect the target and background covariances to line up in this way, but it simplifies the analysis (whitening the background has the effect of whitening the target as well), and is at least more realistic than assuming the target has zero covariance. In other work employing the Gaussian distribution to model the spectral variability of different target materials, each specific material was considered having its own covariance matrix [65].

Tyo *et al.* [66] evaluated several monovariate distributions besides Gaussian for modeling both the single-band spectral radiance and the total radiance (summed across all bands) extracted from multiple instances of a same object in a hyperspectral image. It was observed that the extracted data exhibited a skewed distribution and thus Lognormal, Gamma, and Weibull pdfs were employed for fitting. Experiments showed that none of the tested pdfs did a very good job in accurately modeling the statistical variability of the spectral signature (with a slight advantage of the Lognormal distribution when the total radiance was modeled) and suggested testing other more complex distributions. In [67], the univariate Beta distribution was found effective at modeling the spectral variability affecting target material reflectances at each spectral dimension while incorporating distributional skew and assuring that the reflectance values of the target materials are constrained to a physically realistic range. In order to approximate any distribution that the target materials may exhibit, such as multi-modal distributions very commonly found in real data, a mixture of Gaussian distributions was employed in [68]. These models based on beta and mixture of Gaussian distributions were specifically employed in unmixing applications in the context of endmember spectral variability [67] [68]. Simplex-based distributions have also been suggested for anomaly detection [69].

B. Unconstrained subspace models

Subspaces have been widely utilized to model the spectral variability of observed material spectra [36], [45], [46], [70], [71]. With a subspace model, the target material vector is restricted to vary in a K-dimensional subspace of the d-dimensional data space (with $K \leq d$). The amount of variability allowed increases as K increases from $K = 1$ (scalar variability in one dimension) to $K = d$ (the spectral vector is allowed to vary over all
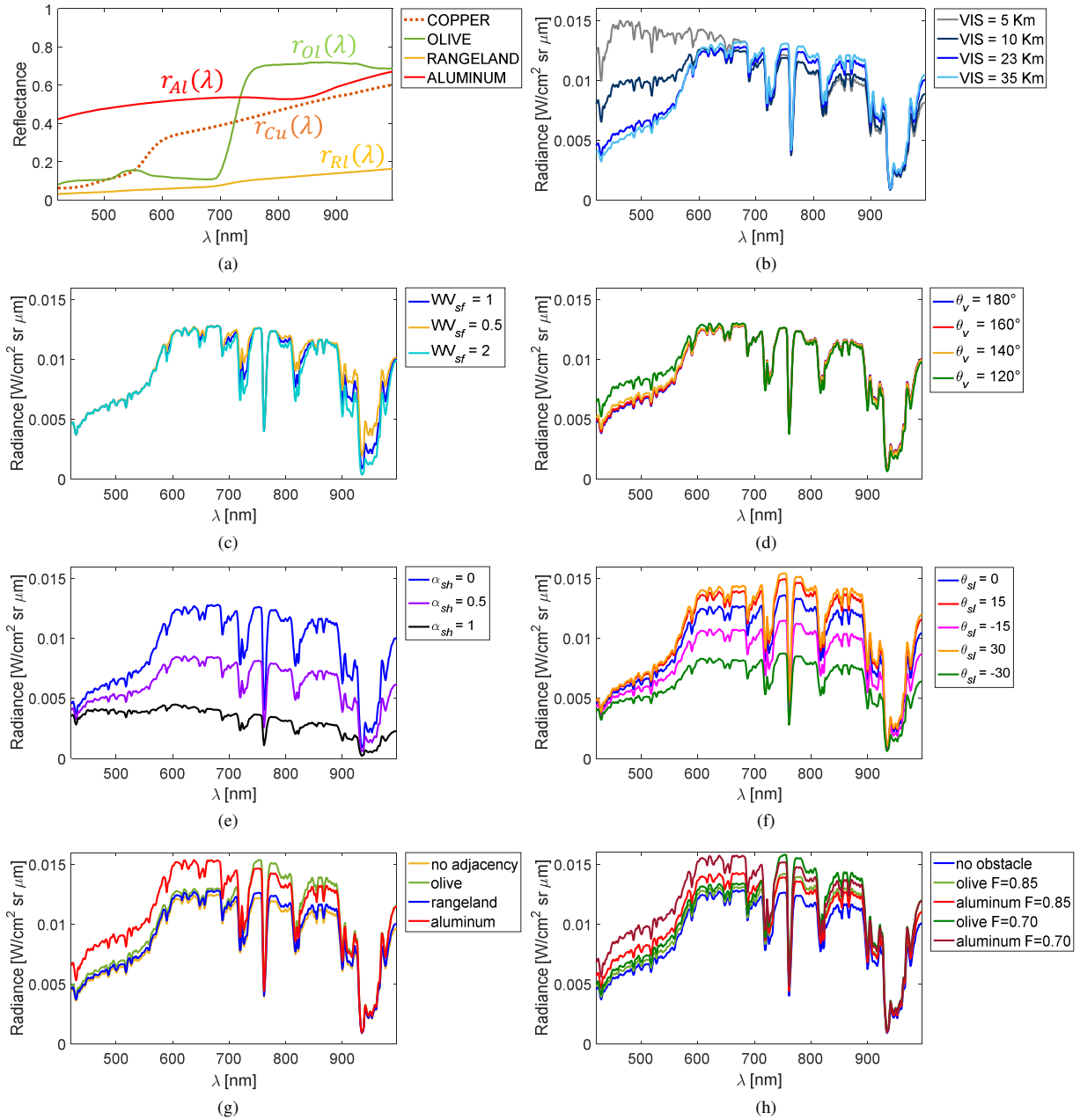
Fig. 4. Example showing the major effects of environmental variability sources over the spectrum of a copper object. (a) Spectral signatures employed. The dotted signature $r_{Cu}(\lambda)$ is the reflectance of the copper object that served as target. The other signatures were used to simulate adjacency effects and secondary illumination by obstructing obstacles. (b) Effects of visibility (VIS). (c) Effects of water vapor scaling factor (WV$_{sf}$). (d) Effect of sensor viewing angle, measured from the zenith ($\theta_v = 180° - \theta_o$; note that $\theta_v = 180°$ corresponds to the nadir view). (e) Shadowing effects. (f) Topographic effects. (g) Adjacency effects. (h) Effects due to secondary illumination by obstacles.
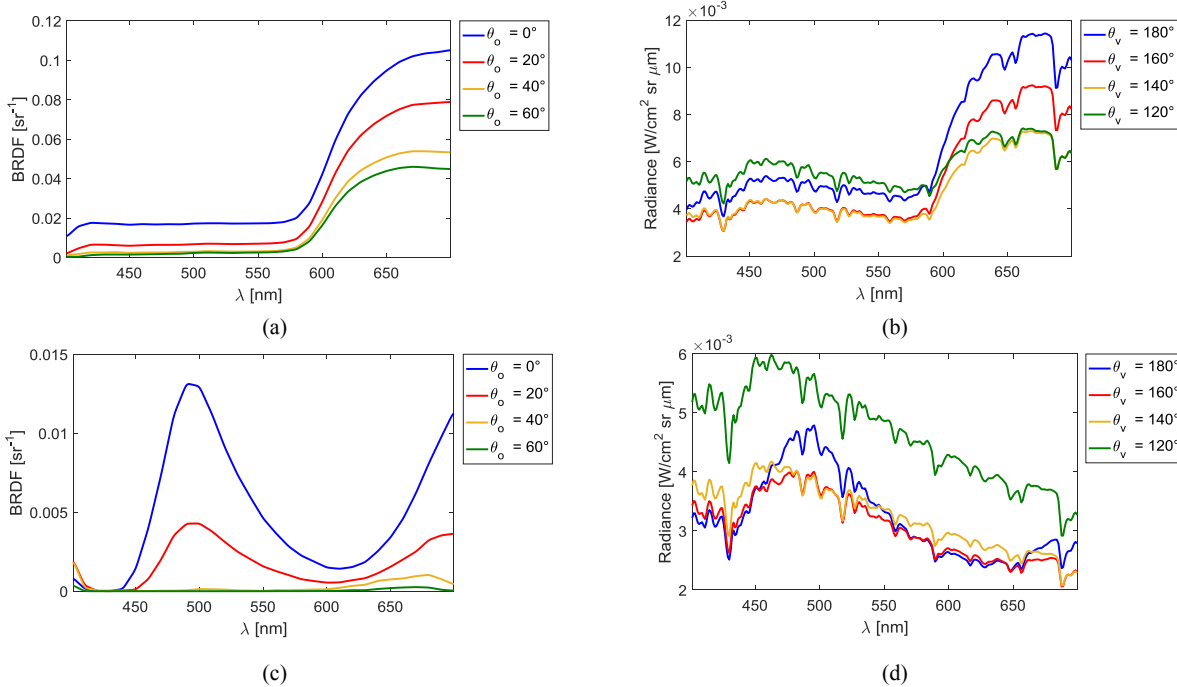
Fig. 5. Example showing the effects of sensor viewing angle variation for two non-Lambertian targets. (a,b) Target painted with 'Garnet red' lacquer. (c,d) Target painted with the glossier 'Mystique' lacquer. (a,c) BRDF of the targets plotted for a fixed illumination angle ($\theta_i = \theta$) and different orientation angles θ_o. In the 'Garnet red' target, BRDF effects mostly consist in scaling of the reflectance functions with varying orientation angles, whereas the 'Mystique' target actually changes color depending on θ_o. (b,d) Corresponding at-sensor radiances for sensor viewing angles $\theta_v = 180° - \theta_o$ matching the orientation angles shown in (a,c).

of $\mathbb{R}^d$). This can be written as:

$$\mathbf{t} = \sum_{k=1}^{K} a_k \mathbf{t}_k = \mathbf{Ta} \tag{13}$$

where $\mathbf{t} \in \mathbb{R}^d$ is the material spectrum, $\mathbf{T}$ is the $d \times K$ target subspace matrix, and $\mathbf{a} \in \mathbb{R}^K$ is a vector of coefficients. According to Eq. (13), the target material spectrum is expressed as a linear combination of the subspace basis vectors $\{\mathbf{t}_k\}_{k=1}^K$ weighted by the elements of the coefficient vector $\mathbf{a}$. If $rank\{\mathbf{T}\} = K$, then the basis vectors are linearly independent, and a given $\mathbf{t}$ will uniquely define the coefficients a_k.

The *affine* model [72] is a useful extension to the subspace model that is obtained by adding a nonzero offset; *i.e.*,

$$\mathbf{t} = \mathbf{t}_o + \mathbf{Ta}. \tag{14}$$

One scenario where this arises is when the coefficient $\mathbf{a}$ corresponds to the amount of direct solar illumination, and $\mathbf{t}_o$ corresponds to the spectrum of the target in the shade.

C. Constrained subspace models

In Eq. (13), the coefficients in the vector $\mathbf{a}$ are unconstrained. If they are to be interpreted as physical quantities (such as abundances), however, then constraints can be used to produce more realistic models of target variability. In particular, non-negativity and sum-to-one constraints for the $\{a_k\}_{k=1}^K$ coefficients can be imposed, thus leading to a simplex model:

$$\begin{cases} \mathbf{t} = \sum_{k=1}^{K} a_k \mathbf{e}_k = \mathbf{Ea} \\ \mathbf{a} \succeq 0 \\ \mathbf{1}^\mathsf{T} \mathbf{a} = 1 \end{cases} \tag{15}$$

where $\mathbf{1}$ is a K-dimensional column vector including all ones, the superscript T denotes the transpose operator, and the columns $\mathbf{e}_k$ of the $d \times K$ matrix $\mathbf{E}$ can be interpreted as the K vertices (or *endmembers*) of the $K - 1$ dimensional simplex, inside of which the target $\mathbf{t}$ is constrained to lie. This is essentially identical the linear mixing model, first proposed by Boardman [73], and now very widely used for background modeling [61], [74]. If the target is being used in an additive model, then the overall magnitude is not important, and there is no reason to enforce the sum-to-one constraint. In this case, the constraining geometry becomes an *uncapped simplex* (*i.e.*, a polyhedral cone with K edges) [71].

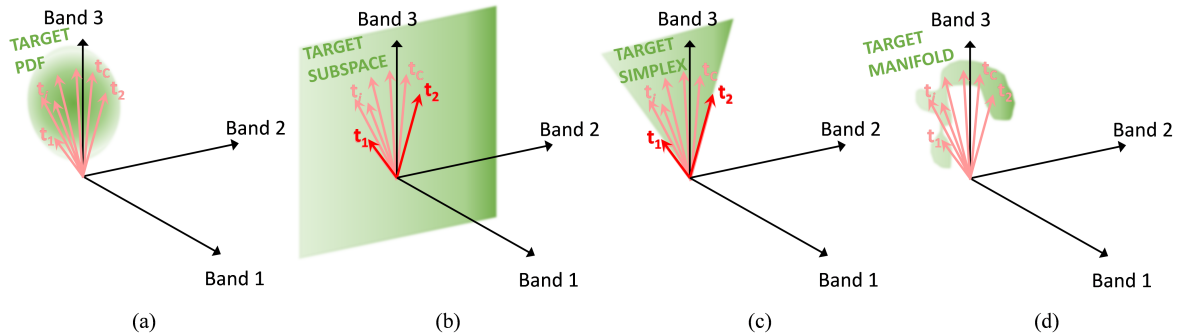With constraints imposed, it is not strictly necessary

Fig. 6. Sketch of target variability models: (a) probabilistic, in which target vectors are drawn from a probability distribution on $\mathbf{t}$, (b) subspace, in which target vectors are linear combinations of $\mathbf{t}_1$ and $\mathbf{t}_2$, (c) constrained subspace, in which target vectors are *positive* linear combinations of $\mathbf{t}_1$ and $\mathbf{t}_2$, and (d) manifold, in which the target vectors lie in a potentially intricate and convoluted structure that locally appears Euclidean.

that the target subspace be of lower dimension than the ambient dimension of the data. The simple geometric approach adopted by Yang *et al.* [75] defines a ball $\mathcal{B}(\mathbf{t}_o, \varepsilon) = \{\mathbf{t} \mid \|\mathbf{t} - \mathbf{t}_o\| \leq \varepsilon\}$ centered on a nominal target spectrum $\mathbf{t}_o$, and having a radius ε corresponding to the variability of the target. The target spectrum is assumed to lie within the ball, and in deriving a detection algorithm, the optimization is constrained by $\mathbf{t} \in \mathcal{B}$.

A different way to constrain a subspace model is by imposing a sparsity constraint on the coefficients. The model becomes a union of lower-dimensional subspaces, and has been employed to produce target detection algorithms (*e.g.*, [76]).

D. Topological manifolds

While the variability of material spectra can sometimes be well-modeled by low-dimensional linear subspaces, the inherent non-linearities of many physical processes that lead to this variability (*e.g.*, that are discussed in Section II) suggest that a more suitable model might be a "curved subspace" or *non-linear manifold* [77], [78]. Manifolds can be linear or non-linear, and *manifold learning* refers to approaches that attempt to recover (*i.e.*, "learn") a lower-dimensional manifold assumed to be embedded in a higher dimensional space [79]. Although hyperspectral data are more commonly modeled with linear manifolds (*i.e.*, constrained and unconstrained subspaces), some hyperspectral data sets have been shown to be more effectively fit with non-linear manifolds [80], [81].

Within this framework, non-linear manifold learning attempts to derive a coordinate system that parameterizes the manifold by, in the words of Bachmann *et al.* [77], "following its intricate and convoluted structure with the hope of achieving a better data representation." In non-linear manifold learning, the concept of linear distance is replaced by that of *geodesic distance* [82],

which corresponds to the length of the shortest path, on the manifold, between two data points. In practice this means that the manifold coordinate system resides on the manifold itself, so that the distances are measured by following the curves of the manifold trajectory and, thus, any geodesic distance along the manifold turns out to be a simple linear distance in the manifold coordinate system [77]. Formally, an m-dimensional manifold $\mathcal{M}$ is defined by stating that for each point $\mathbf{y} \in \mathcal{M}$, there is an open subset (often called a *neighborhood*) $S \subset \mathcal{M}$, with $\mathbf{y} \in S$, that is homeomorphic to an open subset (or neighborhood) S' in an m-dimensional Euclidean space $\mathbb{R}^m$; that is, $S \xrightarrow{g} S'$, where g is the homeomorphic mapping function. Thus, even though a manifold might have a complex non-linear global structure, *locally* it looks like Euclidean space [77].

Well-known manifold learning algorithms include Kernel PCA [83], Isometric Mapping (ISOMAP) [84], Locally Linear Embedding (LLE) [85], [86], and Laplacian Eigenmaps [87]. They have been applied to a variety of hyperspectral image exploitation tasks [77], [80], [81], [88]–[95], but manifold learning in hyperspectral imaging is still a growing research topic [78].

> *"Stay on target."*
> — Gold Five

IV. MANAGING SPECTRAL VARIABILITY IN HYPERSPECTRAL ANALYSIS

As we have seen, spectral variability takes many forms, and there have consequently been many approaches for dealing with this variability in hyperspectral data analysis. In this exposition, we divide those strategies into two categories. The first, which we explore in Section IV-A, is to design algorithms that are robust to this variability. This section will emphasize algorithms

for target detection, but we note that classification, segmentation, and unmixing are tasks that also require attention to spectral variability of materials. The second kind of strategy is to pre-process the data in order to reduce (in some cases, to "project out") the variability due to factors that are not part of our analysis. In this section, we describe both physics-based approaches, such as atmospheric compensation, and more data-driven approaches, such as in-scene target characterization.

Modeling spectral variability of materials is particularly important for target detection. Much of the target detection literature has been focused on how to incorporate variability of the background into the detection algorithms [61], but there remains a need to account for the variability of the target material as well. Even though variability effects for the target material may be milder than those of background, if not properly accounted for they will unavoidably lead to performance degradation. One reason target variability is more difficult to model is the scarceness of labeled training samples, often consisting of a single library spectrum for each material. When, instead, a greater number of labeled training samples are available (as in classification), these may be representative, albeit partially, of the spectral variability to be experienced in the scene. Another difficulty with target detection is the need to transform between reflectance and radiance domains (spectral libraries are in reflectance; remote sensing measurements are in radiance), and this leads to a host of environmental (*e.g.*, atmospheric absorption and scattering, angle of illumination, clouds, shadows, *etc*.) sources of variability.

A. Developing algorithms that are robust to spectral variability

The most mathematically straightforward way to express variability is as an explicit probability density function, and the basic organizing principle for obtaining target detectors is the Likelihood Ratio Test (LRT) [96], which optimally distinguishes between two hypotheses defined by likelihood functions.

For instance, consider the simple case in which a solid (*i.e.*, opaque) target with spectral signature $\mathbf{t}$ covers at least a full pixel. Then the test of whether a given pixel contains a target is of the form:

$$\mathcal{H}_0 : \mathbf{x} = \mathbf{z} \tag{16}$$
$$\mathcal{H}_1 : \mathbf{x} = \mathbf{t} \tag{17}$$

where $\mathbf{x}$ is the measured spectrum at the pixel of interest, and $\mathbf{z}$ represents the background. If we write $p_t(\mathbf{x})$ as the distribution that corresponds to target variability and $p_z(\mathbf{x})$ is the distribution of the background, then the ratio

$$\mathcal{D}(\mathbf{x}) = \frac{p(\mathbf{x}|\mathcal{H}_1)}{p(\mathbf{x}|\mathcal{H}_0)} = \frac{p_t(\mathbf{x})}{p_z(\mathbf{x})} \tag{18}$$

defines an optimal detector for the target in this background. For a given threshold η, we have that $\mathcal{D}(\mathbf{x}) > \eta$ corresponds to the declaration that there is target at this location (alternative hypothesis $\mathcal{H}_1$), whereas $\mathcal{D}(\mathbf{x}) < \eta$ indicates that the target is absent (null hypothesis $\mathcal{H}_0$). Observing that any monotonic function of a detector is an equivalent detector, we could also employ, for instance, $\log \mathcal{D}(\mathbf{x})$. Note that if both $p_t(\mathbf{x})$ and $p_z(\mathbf{x})$ are Gaussian distributions, then the log likelihood detector is in general a quadratic function of $\mathbf{x}$ [97]. If the means are (nearly) equal but the covariances are different, then a Fukunaga-Koontz [98] detector is called for; the Fukunaga-Koontz transform can also be used for dimension reduction for this quadratic detector [99].

If $p_t(\mathbf{x})$ and $p_z(\mathbf{x})$ are Gaussian distributions *with the same covariance*, then the optimal detector is linear – indeed, this is the Fisher linear discriminant [97]:

$$\mathcal{D}(\mathbf{x}) = (\mathbf{t} - \boldsymbol{\mu})^\mathsf{T} R^{-1} (\mathbf{x} - \boldsymbol{\mu}). \tag{19}$$

Here, R is the common covariance, $\mathbf{t}$ is the mean target value, and $\boldsymbol{\mu}$ is the mean background value. Although linear detectors are popular and often effective, the argument above is a poor motivation, because there is little physical reason to expect target and background to have the same covariance.

"An elegant weapon for a more civilized age."
— Obi-Wan Kenobi
(in reference, no doubt,
to clairvoyant fusion)

1) Composite hypothesis testing: When the distinction has to be made between two *families* of hypotheses [100], composite hypothesis testing is required. To illustrate with an example, consider the additive model in Eq. (2) as a hypothesis test:

$$\mathcal{H}_0 : \mathbf{x} = \mathbf{z} \tag{20}$$
$$\mathcal{H}_1 : \mathbf{x} = \mathbf{z} + \epsilon\mathbf{t}, \text{ with } \epsilon \neq 0 \tag{21}$$

The likelihood ratio leading to an optimal detector is given by

$$\mathcal{D}(\epsilon; \mathbf{x}) = \frac{p_z(\mathbf{x} - \epsilon\mathbf{t})}{p_z(\mathbf{x})} \tag{22}$$

but in order to employ this expression, one has to know what value of ϵ to use in the numerator. In general, one does not know ϵ (indeed, if ϵ were actually known, there would be no need to test for whether ϵ were nonzero). For this reason, the detector in Eq. (22) is called a "clairvoyant" detector [96]. The hypothesis that $\epsilon \neq 0$ is a *composite* hypothesis because it encompasses a family of simple hypotheses corresponding to specific values of ϵ. What this means is that the LRT in Eq. (22) is

insufficient; we don't know ϵ so we need a detector $\mathcal{D}(\mathbf{x})$ that does not depend on ϵ.

In what follows, we will describe several strategies for producing an ϵ-independent $\mathcal{D}(\mathbf{x})$, but first we will consider how to evaluate the quality of such a detector. This is not entirely trivial, either, because the performance of $\mathcal{D}(\mathbf{x})$ does depend on ϵ. For any given ϵ, we know the pdf for the target and background classes and from that, we can determine the *power* of a detector (its detection probability at a given false alarm rate) and can tell whether one detector is more powerful than another. If the one detector is more powerful than the other for *all* values of ϵ, then it is said to be *uniformly* more powerful. Further, if there is a detector that is uniformly more powerful than *any* other detector, then it is a *uniformly most powerful* (UMP) detector. The UMP detector is the holy grail of composite hypothesis testing; if you can get it, that is the detector you want. Unfortunately, not all problems admit a UMP detector.

One of the few problems that do admit a UMP detector is the additive target model with a Gaussian background distribution. In this special case, the optimal detector $\mathcal{D}(\epsilon; \mathbf{x})$ defined in Eq. (22) is effectively independent of ϵ. In particular, for any value of ϵ, one can obtain $\mathcal{D}(\epsilon; \mathbf{x})$ from a simple monotonic rescaling of

$$\mathcal{D}(\mathbf{x}) = \mathbf{t}^\mathsf{T} R^{-1}(\mathbf{x} - \boldsymbol{\mu}) \tag{23}$$

which is the Adaptive Matched Filter (AMF) [101]–[103]. The AMF and the Fischer discriminant in Eq. (19) are similar in appearance, differing only in their leading factor – $\mathbf{t}$ for AMF versus $(\mathbf{t} - \boldsymbol{\mu})$ for Fisher – and the Fisher discriminant is sometimes referred to as a "matched filter" (*e.g.*, in [97]). But it is important to remember that they are derived under quite different assumptions.

A more typical case arises when, for instance, $p_z(\mathbf{x})$ is a multivariate t pdf. One can still derive an expression for $\mathcal{D}(\epsilon; \mathbf{x})$ in that case, but the different values of ϵ will be different detectors; they cannot be rescaled into a common ϵ-independent detector.

The most widely used approach for dealing with the composite hypothesis problem is the Generalized LRT (GLRT), in which the variability is parameterized and maximum likelihood estimates of the parameters are sought [96]. In the case of the additive model, the detector becomes $\mathcal{D}(\hat{\epsilon}; \mathbf{x})$, where $\hat{\epsilon}$ is the maximum likelihood estimator:

$$\hat{\epsilon}(\mathbf{x}) = \operatorname{argmax}_\epsilon p_z(\mathbf{x} - \epsilon \mathbf{t}). \tag{24}$$

Thus, the GLRT detector is given by

$$\mathcal{D}(\mathbf{x}) = \mathcal{D}(\hat{\epsilon}(\mathbf{x}); \mathbf{x}) = \frac{\max_\epsilon p_z(\mathbf{x} - \epsilon \mathbf{t})}{p_z(\mathbf{x})}. \tag{25}$$

When the background distribution p_z is Gaussian, this likelihood ratio leads trivially to the AMF. If the background distribution is a fatter tailed elliptically-contoured multivariate t-distribution, one can solve Eq. (24) explicitly, and obtain a closed-form solution for the GLRT detector [104]. In the heavy-tailed limit, this GLRT detector becomes the Adaptive Coherence Estimator (ACE) detector:

$$\mathcal{D}(\mathbf{x}) = \frac{\mathbf{t}^\mathsf{T} R^{-1}(\mathbf{x} - \boldsymbol{\mu})}{\sqrt{(\mathbf{x} - \boldsymbol{\mu})^\mathsf{T} R^{-1}(\mathbf{x} - \boldsymbol{\mu})}}. \tag{26}$$

ACE is a detector that had previously been derived using different assumptions [105], and that nonetheless has enjoyed remarkable success in many situations that are in clear violation of those initial assumptions [106].

If instead of the additive target model we use the replacement target model in Eq. (3), then we cannot have a UMP detector. Here, α is the nuisance parameter, and to obtain a GLRT solution, we must solve

$$\hat{\alpha} = \operatorname{argmax}_\alpha p_x(\mathbf{x}|\alpha)$$
$$= \operatorname{argmax}_\alpha (1 - \alpha)^{-d} p_z\left(\frac{\mathbf{x} - \alpha \mathbf{t}}{1 - \alpha}\right). \tag{27}$$

It turns out that this can be solved in closed form when p_z is Gaussian [107], leading to the finite target matched filter (FTMF) detector. The result has been further extended to a family of elliptically-contoured background distributions [108], [109].

A generalization of the GLRT concept is the *clairvoyant fusion* (CF) approach [110], [111]. In this approach, one begins with a clairvoyant detector; *i.e.*, the detector $\mathcal{D}(\epsilon; \mathbf{x})$ that is optimal if ϵ is the nonzero value in the alternative hypothesis. Then we can express the result in Eq. (25) by writing

$$\mathcal{D}(\mathbf{x}) = \max_\epsilon \mathcal{D}(\epsilon; \mathbf{x}). \tag{28}$$

We observe that this GLRT detector is a "max-fusion" over clairvoyant detectors. Recognize, however, that if $\lambda(\epsilon)$ is any positive function of ϵ, then $\mathcal{D}^*(\epsilon; \mathbf{x}) = \lambda(\epsilon)\mathcal{D}(\epsilon; \mathbf{x})$ is also a clairvoyant detector, since it is just a monotonic rescaling. We can now do a max-fusion over this new family of clairvoyant detectors to create a new CF detector:

$$\mathcal{D}(\mathbf{x}) = \max_\epsilon \mathcal{D}^*(\epsilon; \mathbf{x}) \tag{29}$$
$$= \max_\epsilon \lambda(\epsilon)\mathcal{D}(\epsilon; \mathbf{x}) \tag{30}$$
$$= \frac{\max_\epsilon \lambda(\epsilon) p_z(\mathbf{x} - \epsilon \mathbf{t})}{p_z(\mathbf{x})}. \tag{31}$$

Because the choice of $\lambda(\epsilon)$ is virtually limitless (it only matters that it is positive for all values of ϵ), the CF framework provides great flexibility in deriving a detector. There is unfortunately little guidance with respect to the "best" choice of the function $\lambda(\epsilon)$, but

a clever practitioner may be able to craft a $\lambda(\epsilon)$ that enables the inversion

$$\hat{\epsilon}(\mathbf{x}) = \operatorname{argmax}_{\epsilon} \lambda(\epsilon)\mathcal{D}(\epsilon; \mathbf{x}) \qquad (32)$$

to be analytically tractable, thus leading to a closed-form solution for the detector.

When *prior* probability density functions are available, Bayesian composite hypotheses testing can be performed [100], [112]. Here, instead of taking a maximum over clairvoyant detectors, one takes a weighted average. Again using the additive target model as an example, we have, in contrast to Eq. (25) or Eq. (31):

$$\mathcal{D}(\mathbf{x}) = \frac{\int_{\epsilon} d\epsilon\, \pi(\epsilon)\, p_z(\mathbf{x} - \epsilon\mathbf{t})}{p_z(\mathbf{x})} \qquad (33)$$

where $\pi(\epsilon)$ is the prior; it is an arbitrary non-negative function of ϵ that is chosen by the practitioner.

None of these three approaches – GLRT, CF, or Bayesian – can be counted on to produce a UMP detector, because such detectors very often simply do not exist. But when UMP is not available, a next-most-desirable trait in a detector is that it be *admissible*. A detector is admissible if there does *not* exist another detector that is uniformly more powerful than it is. Although the integral in Eq. (33) often makes the Bayesian detectors less convenient than GLRT or CF detectors that are associated with closed-form expressions, the Bayesian detectors have the following important theoretical advantage: they are provably admissible [100]. By contrast, some GLRT and CF detectors are not admissible [111].

The notion of "admissible" detectors is of great theoretical importance, but it is often remarked that perfectly respectable (*i.e.*, useful) detectors may still be formally inadmissible.

While beyond the scope of this review, a natural follow-on would be to determine the sensitivity of different algorithms to material variability; algorithms that are more "sophisticated" (*i.e.*, more finely-tuned to optimizing performance with respect to a special case) may be less robust to deviations from those special cases. That is, the simpler algorithms may in practice perform better.

2) Invariant approaches: By imposing the constraint on a target detector that it be explicitly invariant to some aspect of that target's variability, we thereby obtain an effective robustness to that variability. This means that we are allowed to have a given amount of ignorance of the variables that alter the target spectrum amplitude and shape [113]. Formally, if we consider that the target spectrum $\mathbf{t}$ depends on some variable parameters enclosed in the vector $\boldsymbol{\Theta}$, an invariant detector should ideally produce the same output (in terms of declaring

$\mathcal{H}_0$ or $\mathcal{H}_1$) regardless of $\boldsymbol{\Theta}$. This condition is assured only in some special cases (see below) [114]. In many other cases, the condition is relaxed to obtain a detector that is robust (though not strictly invariant) to target variability. In striving toward obtaining invariant detectors, researchers have invoked a variety of models that are described in Section III.

Through the exploitation of probabilistic models (see Section III-A), GLRT invariance has been widely investigated [100], [113]–[116]. The GLRT has been found to be invariant (in terms of the condition stated above) with respect to a group of transformations if, upon transformation, the pdfs conditioned to the two hypotheses remain in the same family and the parameter spaces are preserved [100], [114]. Formally, we can write [114]:

$$\begin{cases} f \in \mathcal{F} \\ p(\mathbf{x}; \boldsymbol{\Theta}) = p(f\mathbf{x}; f\boldsymbol{\Theta}) \\ f : \Omega \to \Omega \end{cases} \implies \mathcal{D}(f\mathbf{x}) = \mathcal{D}(\mathbf{x})$$

$$(34)$$

where $\boldsymbol{\Theta}$ is the parameter vector that belongs to the parameter spaces Ω_0 and Ω_1 under the null and alternative hypotheses, respectively, and $f \in \mathcal{F}$ is the set of transformations with respect to which the GLRT is invariant. A proof is given in [114]. In [115], the relationships between the GRLT and the UMP invariant (UMPI) tests are described. Other approaches to find UMPI tests are illustrated in [113].

Aside from GLRT and UMPI tests, the most widely employed invariant approaches are invariant in a broader sense, in that they are robust to variability (*i.e.*, their performances are not as degraded as conventional detectors), while not strictly assuring the same detector output upon variation of the target signature. Nonetheless, in the literature they are called "invariant methods" and we will adhere to this nomenclature.

One straightforward way to impose invariance is by invoking (unconstrained) subspace models (see Section III-B). If $\mathcal{T}$ is a subspace, then it can be parameterized by a matrix $\mathbf{M}$, so that $\mathbf{t} \in \mathcal{T}$ is equivalent to the existence of a vector of coefficients $\mathbf{p}$ with $\mathbf{t} = \mathbf{Mp}$.

One general approach for finding this target subspace is to obtain (by measurement or simulation or both) a discrete (and ideally large) set $\mathcal{T} = \left\{\mathbf{t}^{(i)}\right\}_{i=1}^{C}$ of radiance spectra for a given target material. This set should span the range of variation over which robustness is desired, and can include intrinsic, extrinsic, and/or environmental variability [13], [35], [36], [45]–[47], [117]. For example, Fig. 4 illustrates spectral variability under a range of atmospheric conditions by producing such an ensemble of individual spectra.

The idea of using a low-dimensional subspace is that for some $Q \ll C$, we can find a model characterized by a $Q \times d$ matrix $\mathbf{M}$ that spans a subspace that approximates

each of the target spectra in the set, *i.e.*,

$$\mathbf{t}^{(i)} \approx \sum_{k=1}^{Q} p_k^{(i)} \mathbf{m}_k = \mathbf{M}\mathbf{p}^{(i)}. \qquad (35)$$

From the subspace defined by $\mathbf{M}$ we can invoke, for instance, subspace versions of AMF or ACE [97], [116]. This approach has been extensively pursued in the literature [13], [35], [36], [45]–[47], [71], [117], [118]. Here, the set of target radiance spectra have generally been synthetically generated by radiative transfer modeling and variability has been incorporated by varying the environmental, acquisition geometry, and illumination parameters, together with variations of the BRDF by exploiting suitable physics-based models [13], and in some cases also exploiting in-scene information [117] or large spectral library databases [3].

It should be said that having to identify a low-dimensional linear subspace poses the perennial challenge of estimating the "optimal" subspace dimensionality. If the dimension is too small, then the approximation in Eq. (35) becomes poor; but as the dimension increases, the subspaces can harbor ever larger numbers of false alarms [71]. Another general drawback of this approach is that, regardless of its dimensionality, the size of the subspace spanned by the basis vectors is not constrained because the coefficients $p_k^{(i)}$ can take any (even unphysical) values [119].

One way to deal with subspace high dimensionality issues is to constrain the subspace by imposing conditions on the coefficients $p_k^{(i)}$, as for instance is done with constrained subspace or simplex modeling (see Section III-C). Various approaches have been explored for simplex-based variable target detection [71], [119], [120], *e.g.* by enforcing only the additivity constraint [71], by enforcing both constraints [119], or by performing simultaneous linear unmixing on the target and background spectra together to obtain physical abundances [120]. This has led to methods such as Simplex ACE or Simplex AMF [71], which have exhibited detection performances that, in contrast with their unconstrained subspace-based counterparts, are stable with respect to increasing amounts of spectral variability and increasing sizes of the target subspace.

More recent approaches to building invariant algorithms resort to non-linear models, such as manifolds (see Section III-D). As previously noted, not all variability effects on spectra can be adequately captured by employing low-dimensional *linear* models, but the variable target spectra often "live" within a lower-dimensional non-linear manifold. In [89], the target radiance spectra $\left\{\mathbf{t}^{(i)}\right\}_{i=1}^{C}$ are assumed to lie within a manifold, and a graph-based manifold learning approach is employed to perform invariant target detection. Resorting to non-

linear manifolds is often necessary in underwater hyperspectral remote sensing as well – where water spectra change non-linearly upon variations of depth and water inherent optical properties [81]. Non-linear dimensionality reduction by manifold learning has been successfully exploited in several hyperspectral underwater remote sensing applications, such as underwater object detection [81] and hyperspectral bathymetry [90].

One challenge with using radiative transfer modeling to create a very large set of variable target spectra is that the computation can be prohibitive. It is worth mentioning that some "emulation" approaches have been proposed, in which a reduced number of available variable target spectra is used to learn a stochastic model function of a given number of predictor variables. The learned model is then employed to generate an arbitrary number of variable target spectra with potential real-time applications [121].

3) Matched-pair machine learning approach: As is evident in, for instance, the additive and replacement models, the observed spectrum at a given pixel is often a combination of target and background. As a consequence. target variability is often mixed up with background variability. Further, while target variability can be informed by *ab initio* physics modeling, the background variability is best estimated by looking at actual background pixels. The aim of the matched-pair machine learning (MPML) framework [109], [122] is to produce a target detection algorithm that is customized to the observed background while at the same time exploiting a model of target variability and how the target and background interact. In MPML, every background pixel $\mathbf{z}$ is paired with a background-plus-target pixel $\mathbf{x} = \mathbf{z} + \epsilon \mathbf{t}$ (in the case of the additive target model), and a training set of these matched pairs is used to construct a machine learning classifier to distinguish background from background-plus-target. Spectrally-variable targets can be dealt with very naturally in this framework; the $\mathbf{t}$ that is used to create the target-present half of the matched pair is drawn from a distribution $p_t(\mathbf{t})$ that characterizes target spectral variability.

In fact, MPML can be run in a "transductive" mode [123], in which the classifier is trained on the same data to which it is applied. Normally this is problematic, but it works here because MPML does not actually use labeled data. When the classifier is applied to the original hyperspectral image pixels, it will identify those pixels that are most like the artificial background plus target pixels in the training set; that is, it will find pixels in the original image that are likely to have target in them. There is an implicit assumption that the vast majority of the pixels in the original image are target-free background pixels, so the contamination effect of the

few target-containing pixels in the training step will be minimal. It is important that the classifier is not overfit to the data (and cross-validation is useful or assessing that); an overfit classifier will classify all the original pixels as target-free, even those few that do have target.

4) Other machine learning approaches: It is expected that machine learning approaches, and non-parametric fitting in general, will continue to garner attention as tools for modeling variability. Hyperspectral target detection is enhanced by using these tools for characterizing the background in a local [124] or context dependent [125], [126] manner.

5) Robustness to variability in other hyperspectral image analysis tasks: Although we have focused here on robustness to variability for target detection, many efforts can be found throughout the hyperspectral image analysis literature for coping with spectral variability of materials. Endmember extraction and unmixing methods, for instance, in their simplest forms, assume "pure" endmembers, but for good performance on more realistic hyperspectral data, algorithms require robustness to variability within the endmember classes [74], [127], [128]. Clustering and material discrimination methods have also been developed to handle intra-class spectral variability [128], [129]. The variability due to noise and bottom reflectance materials has been taken into account in retrievals of sub-surface reflectance in shallow water remote sensing [130]. Change detection algorithms require suppressing the pervasive differences between images (due to environmental effects, sensor calibrations, shadows, *etc.*) in order to identify the interesting changes [131]–[133].

B. Mitigating spectral variability in hyperspectral data

Instead of designing exploitation methods that are robust to variability, sometimes it may be worth trying to remove or mitigate the variability prior to proceeding with the hyperspectral image exploitation task. In the literature, this approach has mostly been pursued to back out environmentally-induced variability such as atmospheric and illumination effects [134], shadows [32], [135], and clouds [136], but approaches aimed at mitigating variability due to material morphology and composition have also been proposed [137], [138]. In the following, we provide a brief overview of these approaches.

1) Atmospheric compensation: The most widely employed pre-processing step aimed at removing the environmentally-induced variability is undoubtedly Atmospheric Compensation (AC). Literally, AC refers to the procedure of obtaining an approximate ground-leaving reflectance image starting from the calibrated spectral radiance hyperspectral data cube [134]. This approach is generally preferred by image analysts such as geologists or spectroscopists, although in some applications where a target material needs to be searched for within a hyperspectral image, the so-called forward modeling AC approach of converting the spectral library reflectance into an at-sensor radiance spectrum may be pursued [36]. AC methods can be broadly divided into two categories: empirical scene-based methods [139]–[144] and radiative transfer modeling based methods [38], [134], [145]–[147], although some hybrid approaches may be found as well [148]. Empirical methods exploit data directly from the radiance image such as the average spectrum of the scene [139], or the spectrum of a "spectrally flat" material within the scene [143], or vegetation and shade spectra [149]. Among these, the popular Empirical Line Method (ELM) [142], [144] requires filed measurements of reflectance spectra for at least one dark and one bright reference material and linearly regresses the imagery radiance data against the reflectance field measurements; this enables the derivation of gain and offset curves that are then applied to each image pixel. Like most empirical methods, ELM does not require absolute radiometric calibration, but does require further adjustments if the atmospheric and illumination conditions are not stationary within the scene [150].

Whereas empirical methods are widely applied in operational scenarios [151], over the years researchers have put their efforts towards "forever improving the physical model" [152] to provide increasingly more accurate radiative transfer modeling based AC techniques. The AC approaches belonging to this category are based on an analytic model for the radiation transfer in the atmosphere such as that explored in Section II-C, and thus need to estimate the various model parameters and then invert the model to retrieve the surface reflectance. Model parameter estimation is generally accomplished by (i) resorting to radiative transfer codes such as MOD-TRAN [54] or 5S [153], or (ii) exploiting the imaged radiance in certain bands to estimate visibility, spatially variable parameters (*e.g.*, water vapor and aerosols), and adjacency effects. One of the first AC techniques led to the atmospheric removal program (ATREM) [145], which did not explicitly account for adjacency effects. This was instead incorporated in the approach Fast Line-of-sight Atmospheric Analysis of Spectral Hypercubes (FLAASH) [154], which is a very popular commercial AC software package. More recently, the Atmospheric and Topographic Correction algorithm (ATCOR) [147] has been developed which also features a correction for topographic effects (provided that a digital elevation map

[DEM] of the scene is available).

2) Shadow and cloud removal: Shadows and clouds are among the most non-stationary sources of variability in hyperspectral imagery, both in space and in time. Whereas the effects of clouds are mostly significant for high-altitude acquisitions (*e.g.*, from spaceborne sensors), the impacts of shadows are stronger for high spatial resolution imagery, especially in the urban environment.

Approaches to shadow removal are typically based on specific transformations that are applied to the imaged data [32], [155]. For instance, a transformation from Cartesian to hyperspherical coordinates allows the data to be separated into $d-1$ angular components related to spectral information content, and one radial component whose magnitude represents the overall brightness level [32]. By segmenting the image and identifying those clusters with the lowest magnitude of the radial component, a shadow mask can be obtained [32]. Another approach to remove shadows is to exploit ancillary data, such as light detection and ranging (LIDAR) 3D point clouds [135] [44].

For clouds, their detection is essentially based on finding pixels in which the cloud optical depth is greater than a specific threshold [136], or evaluating a combination of threshold tests [156]. This enables generation of a cloud mask. When possible, the radiative effects (*e.g.*, changes in illumination, shadows, temperature variations, spatial structure, enhanced adjacency effects) of cloud layers are corrected [136].

3) Spectral reflectance decomposition with multiple Gaussian curves: A physics-inspired spectral characterization for material reflectance spectra was presented by Lanker *et al.* [137], with the aim of suppressing variability due primarily to material morphology and composition. Inspiration was drawn from the Lorentz oscillator model applied to the material dielectric function (*i.e.*, how, as a function of wavelength, the optical index of refraction varies). Each Lorenz oscillator, roughly speaking, corresponds to a spectral "line" (though for solid materials, unlike gases, individual lines can be very broad); thus, a model of material spectral reflectance with multiple Lorentz oscillators identifies multiple lines in a spectrum. In [137], the authors choose to fit the curve of reflection versus wavelength using Gaussian line shapes – this was found to be simpler and less computationally intense than fitting the complex dielectric function with Lorentz oscillators). This is similar to the curve fitting approaches performed in mineral spectroscopy [157].

Indeed, it was experimentally found that whereas the Gaussian amplitudes were subject to considerable variations upon varying atmospheric, illumination, and viewing conditions as well as material morphology and composition, the locations and widths of the Gaussian functions were much less affected [137], [138]. These findings suggest that alternative profiles based on locations and widths of the Gaussian functions used to fit the spectra could be adopted for robust characterization of target material spectra, and potentially for background spectral variability as well.

4) In-scene target signature characterization: In some circumstances, the problem of target variability can be bypassed by directly estimating from the image the most discriminative target spectral signature for a given scenario and material by means of a multiple-instance machine learning approach [158]. This can be accomplished when at least imprecisely-labeled training samples are available - for instance, when Global Positioning System (GPS) coordinates for the targets are known and general regions of pixels containing the targets can be identified (GPS precision is generally limited by the co-registration accuracy and thus pixel-level ground truthing is hardly ever achieved). With a simple iterative algorithm with a closed-form solution for the update rule, the Multiple Instance Learning (MIL) approach is capable of retrieving the target signature that maximizes the cosine spectral similarity (more robust in case of mixed training data) between the estimated signature and the positive (target-including) instances, while minimizing the similarity with the non-target labeled instances. As a byproduct, the estimated signature helps in uncovering the discriminative spectral characteristics and features of the target class, otherwise deeply buried within the mixed and imprecisely-labeled training pixels [158].

> *"Only a Sith deals in absolutes."*
> — Obi-Wan Kenobi

V. Conclusions

Spectral variability fundamentally underlies the whole problem of analyzing, understanding, and interpreting hyperspectral imagery.

Of course, spectral variability is not all bad. It is the variability between different materials that allows us to distinguish them in spectral remote sensing. In this discussion, we have emphasized designing algorithms that are *robust* to variability, but for some applications we might want algorithms that are *sensitive* to variability, even within the same material. For instance, in our search for the perfect beach, we may want to find sand of a very particular type, not just sand as a single class.

Our aim in analyzing hyperspectral imagery is to enhance the distinctions we care about (natural vs. man-made, healthy vegetation vs. senescent, target vs. background, *etc.*) while suppressing the distinctions that needlessly confound our analysis (sunlit vs. shadow, humid atmosphere vs. dry, nadir vs. slant angle view, *etc.*), with the realization that *what matters* and *what distracts* both depend on the application at hand.

In this survey, we have discussed many of the causes of spectral variability that complicate algorithms for hyperspectral analysis, we have described some of the models used to characterize that variability, and we have reviewed a number of strategies for dealing with this variability. In many cases, those strategies are a direct consequence of having an explicit model for the variability.

The spectral variability of materials is a topic of relevance for many aspects of hyperspectral image analysis. It arises in land-cover classification, crop health characterization, image segmentation, endmember extraction, spectral unmixing, change analysis, and – of particular interest to the authors of this paper – target detection. We argue that many of the approaches for dealing with spectral variability in target detection carry over to other problems, though their use in target detection is often more straightforward and illustrative.

> *"Sorry about the mess."*
> — Han Solo

References

[1] D. Manolakis, R. Lockwood, and T. Cooley, *Hyperspectral Imaging Remote Sensing: Physics, Sensors, and Algorithms*, Cambridge University Press, 2016.

[2] K. Z. Doctor, C. M. Bachmann, D. J. Gray, M. J. Montes, and R. A. Fusina, "Wavelength dependence of the bidirectional reflectance distribution function (BRDF) of beach sands," *Applied Optics*, vol. 54, pp. F243–F255, 2015.

[3] R.F. Kokaly, R.N. Clark, G.A. Swayze, K.E. Livo, T.M. Hoefen, N.C. Pearson, R.A. Wise, W.M. Benzel, H.A. Lowers, R.L. Driscoll, and A.J. Klein, "USGS spectral library version 7," Data Series 1035, U.S. Geological Survey, 2017.

[4] J. F. Mustard and J. E. Hays, "Effects of hyperfine particles on reflectance spectra from 0.3 to 25 μm," *Icarus*, vol. 125, pp. 145–163, 1997.

[5] T. A. Blake, C. S. Brauer, M. R. Kelly-Gorham, S. D. Burton, M. Bliss, T. L. Myers, T. J. Johnson, and T. E. Tiwald, "Measurement of the infrared optical constants for spectral modeling: n and k values for $(NH_4)_2SO_4$ via single-angle reflectance and ellipsometric methods," *Proc. SPIE*, vol. 10198, pp. 101980J, 2017.

[6] M. I. Mishchenko, G. Videen, V. A. Babenko, N. G. Khlebtsov, and T. Wriedt, "T-matrix theory of electromagnetic scattering by particles and its applications: a comprehensive reference database," *J. Quantitative Spectroscopic Radiation Transfer*, vol. 88, pp. 357–406, 2004.

[7] T. L. Myers, C. S. Brauer, Y.-F. Su, T. A. Blake, R. G. Tonkyn, A. B. Ertel, T. J. Johnson, and R. L. Richardson, "Quantitative reflectance spectra of solid powders as a function of particle size," *Applied Optics*, vol. 54, no. 15, pp. 4863–4875, 2015.

[8] T. N. Beiswenger, T. L. Myers, C. S. Brauer, Y.-F. Su, T. A. Blake, A. B. Ertel, R. G. Tonkyn, J. E. Szecsody, T. J. Johnson, M. O. Smith, and C. L. Lanker, "Experimental effects on IR reflectance spectra: particle size and morphology," *Proc. SPIE*, vol. 9840, pp. 98400I, 2016.

[9] T. A. Reichardt and T. J. Kulp, "Radiative transfer modeling of surface chemical deposits," *Proc. SPIE*, vol. 9840, pp. 98400M, 2016.

[10] T. J. Gibbs and D. W. Messinger, "Contaminant mass estimation of powder contaminated surfaces," *Proc. SPIE*, vol. 10198, pp. 101980M, 2017.

[11] T. L. Myers, R. G. Tonkyn, T. O. Danby, M. S. Taubman, B. E. Bernacki, J. C. Birnbaum, S. W. Sharpe, and T. J. Johnson, "Accurate measurements of the optical constants n and k for a series of 57 inorganic and organic liquids for optical modeling and detection," *Applied Spectroscopy*, vol. 72, no. 4, pp. 535–550, 2018.

[12] T. A. Blake, T. J. Johnson, R. G. Tonkyn, B. M. Forland, T. L. Myers, C. S. Brauer, Y.-F. Su, B. E. Bernacki, L. H. Hanssen, and G. Gonzalez, "Methods for quantitative infrared directional-hemispherical and diffuse reflectance measurements using an FTIR and a commercial integrating sphere," *Applied Optics*, vol. 57, no. 3, pp. 432–446, 2018.

[13] E. J. Ientilucci and M. Gartley, "Impact of BRDF on physics based modeling as applied to target detection in hyperspectral imagery," *Proc. SPIE*, vol. 7334, 2009.

[14] S. Adler-Golden, D. Less, X. Jin, and P. Rynes, "Modeling and analysis of LWIR signature variability associated with 3D and BRDF effects," *Proc. SPIE*, vol. 9840, pp. 98400P, 2016.

[15] J. Svejkosky, E. Ientilucci, S. Richtsmeier, M. Parente, and C. Bachmann, "A hyperspectral vehicle BRDF sampling experiment," *Proc. SPIE*, vol. 9840, pp. 98401D, 2016.

[16] J. R. Schott, *Remote sensing: the image chain approach*, Oxford University Press, USA, 2007.

[17] J. D. Harms, C. M. Bachmann, J. W. Faulring, and A. J. Ruiz Torres, "A next generation field-portable goniometer system," *Proc. SPIE*, vol. 9840, pp. 98400J, 2016.

[18] J. D. Harms, C. M. Bachmann, B. L. Ambeau, J. W. Faul ring, A. J. Ruiz Torres, G. Badura, and E. Myers, "Fully automated laboratory and field-portable goniometer used for performing accurate and precise multiangular reflectance measurements," *J. Applied Remote Sensing*, vol. 11, no. 4, pp. 1–15, 2017.

[19] R. Montes and C. Urena, "An overview of BRDF models," Tech. Rep. LSI-2012-001, University of Granada, 2012.

[20] R. A. Schowengerdt, *Remote Sensing: Models and Methods for Image Processing*, Elsevier: Academic Press, Burlington,MA, USA, 3rd edition, 2007.

[21] B. Hapke, "Bidirectional reflectance spectroscopy: 1. theory," *J. Geophysical Research: Solid Earth*, vol. 86, no. B4, pp. 3039–3054, 1981.

[22] B. Hapke, *Theory of Reflectance and Emittance Spectroscopy*, Cambridge University Press, 2 edition, 2012.

[23] J. Broadwater and A. Banerjee, "A comparison of kernel functions for intimate mixture models," *Proc. 1st IEEE Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, 2009.

[24] J. Broadwater, A. Banerjee, and P. Burlina, "Kernel methods for unmixing hyperspectral imagery," in *Kernel Methods for Remote Sensing Data Analysis*, G. Camps-Valls and L. Bruzzone, Eds., pp. 249–270. Wiley Chichester, 2009.

[25] J. Broadwater and A. Banerjee, "A generalized kernel for areal and intimate mixtures," *Proc. 2nd IEEE Workshop on Hy-*

perspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), 2010.

[26] J. M. Nascimento and J. M. Bioucas-Dias, "Unmixing hyperspectral intimate mixtures," *Proc. SPIE*, vol. 7830, pp. 78300C, 2010.

[27] R. Heylen and P. Gader, "Nonlinear spectral unmixing with a linear mixture of intimate mixtures model," *IEEE Geoscience and Remote Sensing Letters*, vol. 11, no. 7, pp. 1195–1199, 2014.

[28] A. Hayden, E. Niple, and B. Boyce, "Determination of trace-gas amounts in plumes by the use of orthogonal digital filtering of thermal-emission spectra," *Applied Optics*, vol. 35, pp. 2802–2809, 1996.

[29] A. Beer, "Bestimmung der absorption des rothen lichts in farbigen flussigketiten," *Ann. Physik*, vol. 162, pp. 78–88, 1852.

[30] M. Diani, M. Moscadelli, and G. Corsini, "Improved alpha residuals for target detection in thermal hyperspectral imaging," *IEEE Geoscience and Remote Sensing Letters*, vol. 15, pp. 779–783, 2018.

[31] N. Wang, H. Wu, F. Nerry, C. Li, and Z.-L. Li, "Temperature and emissivity retrievals from hyperspectral thermal infrared data using linear spectral emissivity constraint," *IEEE Trans. Geoscience and Remote Sensing*, vol. 49, no. 4, pp. 1291–1303, 2011.

[32] E. A. Ashton, B. D. Wemett, R. A. Leathers, and T. V. Downes, "A novel method for illumination suppression in hyperspectral images," *Proc. SPIE*, vol. 6966, pp. 69660C, 2008.

[33] A. V. Mariano and J. M. Grossmann, "Hyperspectral material identification on radiance data using single-atmosphere or multiple-atmosphere modeling," *J. Applied Remote Sensing*, vol. 4, no. 1, pp. 043563, 2010.

[34] L. Guanter, R. Richter, and H. Kaufmann, "On the application of the MODTRAN4 atmospheric radiative transfer code to optical remote sensing," *Int. J. Remote Sensing*, vol. 30, no. 6, pp. 1407–1424, 2009.

[35] T. V. Haavardsholm, T. Skauli, and I. Kåsen, "A physics-based statistical signature model for hyperspectral target detection," in *Proc. IEEE Int. Geoscience and Remote Sensing Symposium (IGARSS)*, 2007, pp. 3198–3201.

[36] S. Matteoli, E. J. Ientilucci, and J. P. Kerekes, "Operational and performance considerations of radiative-trasnfer modeling in hyperspectral target detection," *IEEE Trans. Geoscience and Remote Sensing*, vol. 49, pp. 1343–1355, 2011.

[37] R. Richter, "Correction of satellite imagery over mountainous terrain," *Applied Optics*, vol. 37, no. 18, pp. 4004–4015, 1998.

[38] C. Miesch, L. Poutier, V. Achard, X. Briottet, X. Lenot, and Y. Boucher, "Direct and inverse radiative transfer solutions for visible and near-infrared hyperspectral imagery," *IEEE Trans. Geoscience and Remote Sensing*, vol. 43, no. 7, pp. 1552–1561, 2005.

[39] J. E. Hay, "Calculation of monthly mean solar radiation for horizontal and inclined surfaces," *Solar Energy*, vol. 23, no. 4, pp. 301–307, 1979.

[40] S. Sandmeier and K. I. Itten, "A physically-based model to correct atmospheric and illumination effects in optical satellite data of rugged terrain," *IEEE Trans. Geoscience and Remote Sensing*, vol. 35, no. 3, pp. 708–717, 1997.

[41] D. Gu and A. Gillespie, "Topographic normalization of Landsat TM images of forest based on subpixel sun-canopy-sensor geometry," *Remote Sensing of Environment*, vol. 64, no. 2, pp. 166–175, 1998.

[42] J. R. Dymond, J. D. Shepherd, G. Gu, and A. Gillespie, "Comment on 'Topographic normalization of Landsat TM images of forest based on subpixel sun-canopy-sensor geometry,' by Gu and Gillespie (Remote Sens. Environ. 64:166-175, 1998)," *Remote Sensing of Environment*, vol. 69, no. 2, pp. 194, 1999.

[43] J. Dozier, "Rapid calculation of terrain parameters for radiation modeling from digital elevation data," *IEEE Trans. Geoscience and Remote Sensing*, vol. 28, no. 5, pp. 963–969, 1990.

[44] K. Ewald, E. J. Ientilucci, A. Buswell, and J. Jacobson, "Improved atmospheric retrievals of hyperspectral data using geometric constraints," *Proc. SPIE*, vol. 9611, 2015.

[45] G. Healey and D. Slater, "Models and methods for automated material identification in hyperspectral imagery acquired under unknown illumination and atmospheric conditions," *IEEE Tran. Geoscience and Remote Sensing*, vol. 37, pp. 2707–2717, 1999.

[46] E. J. Ientilucci and J. R. Schott, "Target detection in a structured background environment using an infeasibility metric in an invariant space," *Proc. SPIE*, vol. 5806, pp. 491–502, 2006.

[47] M. A. Kolodner, "Automated target detection system for hyperspectral imaging sensors," *Applied Optics*, vol. 47, pp. F61–F70, 2008.

[48] P. E. Goa, T. Skauli, and I. Kåsen, "Physical subspace models for invariant material identification: subspace composition and detection performance," *Proc. SPIE*, vol. 5573, pp. 203 – 214, 2004.

[49] R. Richter, M. Bachmann, W. Dorigo, and A. Müller, "Influence of the adjacency effect on ground reflectance measurements," *IEEE Geoscience and Remote Sensing Letters*, vol. 3, pp. 565–569, 2006.

[50] P.-H. Suen, G. Healey, and D. Slater, "The impact of viewing geometry on material discriminability in hyperspectral images," *IEEE Trans. Geoscience and Remote Sensing*, vol. 39, no. 7, pp. 1352–1359, 2001.

[51] J. P. Bishoff, D. W. Messinger, and E. J. Ientilucci, "Oblique hyperspectral target detection," *Proc. SPIE*, vol. 7086, pp. 70860O, 2008.

[52] Y. J. Kaufman and B.-C. Gao, "Remote sensing of water vapor in the near IR from EOS/MODIS," *IEEE Trans. Geoscience and Remote Sensing*, vol. 30, no. 5, pp. 871–884, 1992.

[53] A.M. Baldridge, S.J. Hook, C.I. Grove, and G. Rivera, "The ASTER spectral library version 2.0," *Remote Sensing of Environment*, vol. 113, no. 4, pp. 711–715, 2009.

[54] A. Berk, G. P. Anderson, P. K. Acharya, L. S. Bernstein, L. Muratov, J. Lee, M. Fox, S. M. Adler-Golden, J. H. Chetwynd, M. L. Hoke, R. B. Lockwood, J. A. Gardner, T. W. Cooley, and P. E. Lewis, "MODTRAN5: A reformulated atmospheric band model with auxiliary species and practical multiple scattering options," *Proc. SPIE*, vol. 5425, pp. 341–347, 2004.

[55] N. Acito, S. Matteoli, A. Rossi, M. Diani, and G. Corsini, "Hyperspectral airborne 'Viareggio 2013 Trial' data collection for detection algorithm assessment," *IEEE J. Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 9, no. 6, pp. 2365–2376, 2016.

[56] Cornell University Program of Computer Graphics, "Measurement Data," http://www.graphics.cornell.edu/online/measurements/, 2006.

[57] R. L. Cook and K. E. Torrance, "A reflectance model for computer graphics," *ACM Trans. Graphics*, vol. 1, pp. 7–24, Jan. 1982.

[58] J. P. Kerekes and J. E. Baum, "Full-spectrum spectral imaging system analytical model," *IEEE Trans. Geoscience and Remote Sensing*, vol. 43, no. 3, pp. 571–580, 2005.

[59] M. Shimoni, R. Haelterman, and P. Lodewyckx, "Advancing the retrievals of surface emissivity by modelling the spatial distribution of temperature in the thermal hyperspectral scene," *Proc. SPIE*, vol. 9840, pp. 98400O, 2016.

[60] M. T. Eismann, *Hyperspectral Remote Sensing*, SPIE, 2012.

[61] S. Matteoli, M. Diani, and J. Theiler, "An overview background modeling for detection of targets and anomalies in hyperspectral remotely sensed imagery," *IEEE J. Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, pp. 2317–2336, 2014.

[62] K. Kim and G. Shevlyakov, "Why Gaussianity: An attempt to explain this phenomenon," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 102–113, 2008.

[63] R. S. DiPietro, D. G. Manolakis, R. B. Lockwood, T. Cooley, and J. Jacobson, "Performance evaluation of hyperspectral detection algorithms for sub-pixel objects," *Proc. SPIE*, vol. 7695, pp. 76951W, 2010.

[64] A. Schaum, "Continuum fusion solutions for replacement target models in electro-optic detection," *Applied Optics*, vol. 53, pp. C25–C31, 2014.

[65] A. Zare, P. Gader, and G. Casella, "Sampling piecewise convex unmixing and endmember extraction," *IEEE Trans. Geoscience and Remote Sensing*, vol. 51, pp. 1655–1665, 2013.

[66] J. S. Tyo, J. Robertson, J. Wollenbecker, and R. C. Olsen, "Statistics of target spectra in HSI scenes," *Proc. SPIE*, vol. 4132, pp. 306–314, 2000.

[67] X. Du, A. Zare, P. Gader, and D. Dranishnikov, "Spatial and spectral unmixing using the beta compositional model," *IEEE J. Sel. Topics Appl. Earth Observ.*, vol. 7, pp. 1994–2003, 2014.

[68] Y. Zhou, A. Rangarajan, and P. D. Gader, "A gaussian mixture model representation of endmember variability in hyperspectral unmixing," *IEEE Trans. Image Processing*, vol. 27, pp. 2242–2256, 2018.

[69] J. Theiler, "Ellipsoid-simplex hybrid for hyperspectral anomaly detection," *Proc. 3rd IEEE Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, 2011.

[70] J. C. Harsanyi and C.-I. Chang, "Hyperspectral image classification and dimensionality reduction: an orthogonal subspace projection approach," *IEEE Trans. Geoscience and Remote Sensing*, vol. 32, pp. 779–785, 1994.

[71] A. Ziemann and J. Theiler, "Simplex ACE: a constrained subspace detector," *Optical Engineering*, vol. 56, pp. 081808, 2017.

[72] A. Schaum and R. Priest, "The affine matched filter," *Proc. SPIE*, vol. 7334, pp. 733403, 2009.

[73] J. W. Boardman, "Automating spectral unmixing of AVIRIS data using convex geometry concepts," in *Summaries of the Fourth Annual JPL Airborne Geoscience Workshop*, R. O. Green, Ed., 1994, pp. 11–14.

[74] A. Zare and K. C. Ho, "Endmember variability in hyperspectral analysis: Addressing spectral variability during spectral unmixing," *IEEE Signal Processing Magazine*, vol. 31, pp. 95–104, 2014.

[75] S. Yang, Z. Shi, and W. Tang, "Robust hyperspectral image target detection using an inequality constraint," *IEEE Trans. Geoscience and Remote Sensing*, vol. 53, no. 6, pp. 3389–3404, 2015.

[76] Y. Chen, N. M. Nasrabadi, and T. D. Tran, "Sparse representation for target detection in hyperspectral imagery," *IEEE J. Selected Topics in Signal Processing*, vol. 5, pp. 629–640, 2011.

[77] C. M. Bachmann, T. L. Ainsworth, and R. A. Fusina, "Exploiting manifold geometry in hyperspectral imagery," *IEEE Trans. Geoscience and Remote Sensing*, vol. 43, no. 3, pp. 441–454, 2005.

[78] D. Lunga, S. Prasad, M. M. Crawford, and O. Ersoy, "Manifold-learning-based feature extraction for classification of hyperspectral data: A review of advances in manifold learning," *IEEE Signal Processing Magazine*, vol. 31, pp. 55–66, Jan 2014.

[79] Y. Ma and Y. Fu, *Manifold Learning Theory and Applications*, CRC Press, 2011.

[80] D. Gillis, J. Bowles, G.M. Lamela, W.J. Rhea, C.M. Bachmann, M. Montes, and T. Ainsworth, "Manifold learning techniques for the analysis of hyperspectral ocean data," *Proc. SPIE*, vol. 5806, pp. 342–351, 2005.

[81] D. B. Gillis, "A nonlinear modeling framework for the detection of underwater objects in hyperspectral imagery," *Proc. SPIE*, vol. 9840, pp. 98401B, 2016.

[82] J. Bouttier, P. Di Francesco, and E. Guitter, "Geodesic distance in planar graphs," *Nuclear Physics B*, vol. 663, no. 3, pp. 535–567, 2003.

[83] B. Scholkopf, A. Smola, and K.-R. MÃŒller, "Nonlinear component analysis as a kernel eigenvalue problem," *Neural Computation*, vol. 10, no. 5, pp. 1299–1319, 1998.

[84] J. B. Tenenbaum, V. De Silva, and J. C. Langford, "A global geometric framework for nonlinear dimensionality reduction," *Science*, vol. 290, no. 5500, pp. 2319–2323, 2000.

[85] S. T. Roweis and L. K. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Science*, vol. 290, no. 5500, pp. 2323–2326, 2000.

[86] L. K. Saul and S. T. Rowels, "Think globally, fit locally: Unsupervised learning of low dimensional manifolds," *J. Machine Learning Research*, vol. 4, no. 2, pp. 119–155, 2004.

[87] M. Belkin abd P. Niyogi, "Laplacian eigenmaps and spectral techniques for embedding and clustering," *Advances in Neural Information Processing Systems*, vol. 14, 2001.

[88] A. K. Ziemann, J. Theiler, and D. W. Messinger, "Hyperspectral target detection using manifold learning and multiple target spectra," in *Proc. 44th IEEE Applied Imagery Pattern Recognition (AIPR) Workshop*, 2015.

[89] J. A. Albano and D. W. Messinger, "Spectral target detection using a physical model and a manifold learning technique," *Proc. SPIE*, vol. 8743, pp. 874318, 2013.

[90] C. M. Bachmann and T. L. Ainsworth, "Bathymetric retrieval from manifold coordinate representations of hyperspectral imagery," in *Proc. IEEE Int. Geoscience and Remote Sensing Symposium (IGARSS)*, 2007, pp. 1548 –1551.

[91] A. K. Ziemann and D. W. Messinger, "An adaptive locally linear embedding manifold learning approach for hyperspectral target detection," *Proc. SPIE*, vol. 9472, pp. 94720O, 2015.

[92] H. L. Yang and M. M. Crawford, "Domain adaptation with preservation of manifold geometry for hyperspectral image classification," *IEEE J. Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 9, no. 2, pp. 543–555, 2016.

[93] D. Hong, N. Yokoya, and X.X. Zhu, "Learning a robust local manifold representation for hyperspectral dimensionality reduction," *IEEE J. Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 10, no. 6, pp. 2960–2975, 2017.

[94] L. Zhang, L. Zhang, D. Tao, and X. Huang, "Sparse transfer manifold embedding for hyperspectral target detection," *IEEE Trans. Geoscience and Remote Sensing*, vol. 52, no. 2, pp. 1030–1043, 2014.

[95] L. P. Dorado-Munoz and D. W. Messinger, "Initial study of Schroedinger eigenmaps for spectral target detection," *Optical Engineering*, vol. 55, no. 8, 2016.

[96] S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*, vol. II, Prentice Hall, New Jersey, 1998.

[97] D. Manolakis, "Taxonomy of detection algorithms for hyperspectral imaging applications," *Optical Engineering*, vol. 44, pp. 66403, 2005.

[98] K. Fukunaga and W. L. G. Koontz, "Application of the Karhunen-Loève expansion to feature selection and ordering," *IEEE Trans. Computers*, vol. C-19, pp. 311 – 318, 1970.

[99] X. Huo, M. Elad, A. G. Flesia, R. R. Muise, S. R. Stanfill, J. Friedman, B. Popescu, J. Chen, A. Mahalanobis, and D. L. Donoho, "Optimal reduced-rank quadratic classifiers using the Fukunaga-Koontz transform with applications to automated target recognition," *Proc. SPIE*, vol. 5094, pp. 59–72, 2003.

[100] E. L. Lehmann and J. P. Romano, *Testing Statistical Hypotheses*, Springer, New York, 2005.

[101] I. S. Reed, J. D. Mallett, and L. E. Brennan, "Rapid convergence rate in adaptive arrays," *IEEE Trans. Aerospace and Electronic Systems*, vol. 10, pp. 853–863, 1974.

[102] E. J. Kelly, "An adaptive detection algorithm," *IEEE Trans. Aerospace and Electronic Systems*, vol. 22, pp. 115–127, 1986.

[103] F. C. Robey, D. R. Fuhrmann, E. J. Kelly, and R. Nitzberg, "A CFAR adaptive matched filter detector," *IEEE Trans. Aerospace and Electronic Systems*, vol. 28, pp. 208–216, 1992.

[104] J. Theiler and B. R. Foy, "EC-GLRT: Detecting weak plumes in non-Gaussian hyperspectral clutter using an elliptically-contoured generalized likelihood ratio test," in *Proc. IEEE Int. Geoscience and Remote Sensing Symposium (IGARSS)*, 2008, vol. I, pp. 221–224.

[105] S. Kraut, L. L. Scharf, and R. W. Butler, "The Adaptive Coherence Estimator: a uniformly most-powerful-invariant adaptive detection statistic," *IEEE Trans. Signal Processing*, vol. 53, pp. 427–438, 2005.

[106] D. Manolakis, M. Pieper, E. Truslow, T. Cooley, M. Bruegge-man, and S. Lipson, "The remarkable success of adaptive cosine estimator in hyperspectral target detection," *Proc. SPIE*, vol. 8743, pp. 874302, 2013.

[107] A. Schaum and A. Stocker, "Spectrally selective target detection," in *Proc. Int. Symposium on Spectral Sensing Research (ISSSR)*, 1997, pp. 23–30.

[108] J. Theiler, B. Zimmer, and A. Ziemann, "Closed-form detector for solid sub-pixel targets in multivariate t-distributed background clutter," in *Proc. IEEE Int. Geoscience and Remote Sensing Symposium (IGARSS)*, 2018, pp. 2773–2776.

[109] A. Ziemann, M. Kucer, and J. Theiler, "A machine learning approach to hyperspectral detection of solid targets," *Proc. SPIE*, vol. 10644, pp. 1064404, 2018.

[110] A. Schaum, "Continuum fusion: a theory of inference, with applications to hyperspectral detection," *Optics Express*, vol. 18, pp. 8171–8181, 2010.

[111] J. Theiler, "Confusion and clairvoyance: some remarks on the composite hypothesis testing problem," *Proc. SPIE*, vol. 8390, pp. 839003, 2012.

[112] A. Schaum, "Hyperspectral target detection using a Bayesian likelihood ratio test," *Proc. IEEE Aerospace Conference*, vol. 3, pp. 1537–1540, 2002.

[113] A. Schaum, "Invariance concepts in spectral analysis," *Proc. SPIE*, vol. 10198, pp. 101980H, 2017.

[114] S. M. Kay and J. R. Gabriel, "An invariance property of the generalized likelihood ratio test," *IEEE Signal Processing Letters*, vol. 10, no. 12, pp. 352–355, 2003.

[115] J. R. Gabriel and S. M. Kay, "On the relationship between the GLRT and UMPI tests for the detection of signals with unknown parameters," *IEEE Trans. Signal Processing*, vol. 53, no. 11, pp. 4194–4203, 2005.

[116] L. L. Scharf and B. Friedlander, "Matched subspace detectors," *IEEE Trans. Signal Processing*, vol. 42, no. 8, pp. 2146–2157, 1994.

[117] M. Axelsson, O. Friman, T. V. Haavardsholm, and I. Renhorn, "Target detection in hyperspectral imagery using forward modeling and in-scene information," *ISPRS J. Photogrammetry and Remote Sensing*, vol. 119, pp. 124–134, 2016.

[118] B. Thai and G. Healey, "Invariant subpixel material detection in hyperspectral imagery," *IEEE Trans. Geoscience and Remote Sensing*, vol. 40, no. 3, pp. 599–608, 2002.

[119] S. Adler-Golden, J. Gruninger, and R. Sundberg, "Hyperspectral detection and identification with constrained target subspaces," in *Proc. IEEE Int. Geoscience and Remote Sensing Symposium (IGARSS)*, 2008, vol. II, pp. 465–468.

[120] J. Broadwater and R. Chellappa, "Hybrid detectors for subpixel targets," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 29, pp. 1891–1903, 2007.

[121] E. J. Ientilucci and P. Bajorski, "Stochastic modeling of physically derived signature spaces," *J. Applied Remote Sensing*, vol. 2, no. 1, pp. 023532, 2008.

[122] J. Theiler, "Matched-pair machine learning," *Technometrics*, vol. 55, pp. 536–547, 2013.

[123] J. Theiler, "Transductive and matched-pair machine learning for difficult target detection problems," *Proc. SPIE*, vol. 9088, pp. 90880E, 2014.

[124] J. Theiler and B. Wohlberg, "Regression framework for background estimation in remote sensing imagery," *Proc. 5th IEEE Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, 2013.

[125] H. Jenzri, H. Frigui, and P. Gader, "Context dependent hyperspectral subpixel target detection," *Proc. IEEE Int. Conf. on Image Processing (ICIP)*, pp. 5062–5066, 2014.

[126] T. Glenn, "Context-dependent detection in hyperspectral imagery," *Diss. University of Florida*, 2013.

[127] B. Somers, G. P. Asner, L. Tits, and P. Coppin, "Endmember variability in spectral mixture analysis: A review," *Remote Sensing of Environment*, vol. 115, pp. 1603–1616, 2011.

[128] T. Uezato, R. J. Murphy, A. Melkumyan, and A. Chlingaryan, "A novel endmember bundle extraction and clustering approach for capturing spectral variability within endmember classes," *IEEE Trans. Geoscience and Remote Sensing*, vol. 54, pp. 6712–6731, 2016.

[129] C.-I. Chang, "An information-theoretic approach to spectral variability, similarity, and discrimination for hyperspectral image analysis," *IEEE Trans. Information Theory*, vol. 46, pp. 1927–1932, 2000.

[130] S. Jay, M. Guillame, A. Minghelli, Y. Deville, M. Chami, B. Lafrance, and V. Serfaty, "Hyperspectral remote sensing of shallow waters: Considering environmental noise and bottom intra-class variability for modeling and inversion of water reflectance," *Remote Sensing of Environment*, vol. 200, pp. 352–367, 2017.

[131] A. Schaum and A. Stocker, "Long-interval chronochrome target detection," *Proc. ISSSR (International Symposium on Spectral Sensing Research)*, 1998.

[132] M. T. Eismann, J. Meola, and R. C. Hardie, "Hyperspectral change detection in the presence of diurnal and seasonal variations," *IEEE Trans. Geoscience and Remote Sensing*, vol. 46, no. 1, pp. 237–249, 2008.

[133] J. Theiler, "Quantitative comparison of quadratic covariance-based anomalous change detectors," *Applied Optics*, vol. 47, pp. F12–F26, 2008.

[134] B.-C. Gao, M. J. Montes, C. O. Davis, and A. F. H. Goetz, "Atmospheric correction algorithms for hyperspectral remote sensing data of land and ocean," *Remote Sensing of Environment*, vol. 113, no. SUPPL. 1, pp. S17–S24, 2009.

[135] G. Tolt, M. Shimoni, and J. Ahlberg, "A shadow detection method for remote sensing images using VHR hyperspectral and LIDAR data," in *Proc. IEEE Int. Geoscience and Remote Sensing Symposium (IGARSS)*, 2011, pp. 4423–4426.

[136] J. G. Shanks and B. V. Shetler, "Confronting clouds: Detection, remediation and simulation approaches for hyperspectral remote sensing systems," in *Proc. 29th IEEE Applied Imagery Pattern Recognition (AIPR) Workshop*, 2000, pp. 25–31.

[137] C. L. Lanker and M. O. Smith, "Identification of solid materials using HSI spectral oscillators," *Proc. SPIE*, vol. 9840, pp. 98400U, 2016.

[138] C. L. Lanker and M. O. Smith, "Enhanced detection of solids from Gaussian spectral features," in *Proc. IEEE Int. Geoscience and Remote Sensing Symposium (IGARSS)*, 2017, pp. 1340–1343.

[139] F.A. Kruse, "Use of airborne imaging spectrometer data to map minerals associated with hydrothermally altered rocks in the northern grapevine mountains, Nevada, and California," *Remote Sensing of Environment*, vol. 24, no. 1, pp. 31–51, 1988.

[140] A. M. Filippi, K. L. Carder, and C. O. Davis, "Vicarious calibration of the ocean phills hyperspectral sensor using a coastal tree-shadow method," *Geophysical Research Letters*, vol. 33, no. 22, pp. L22605, 2006.

[141] L. S. Bernstein, X. Jin, B. Gregor, and S. M. Adler-Golden, "Quick atmospheric correction code: Algorithm description and recent upgrades," *Optical Engineering*, vol. 51, no. 11, pp. 111719, 2012.

[142] G. M. Smith and E. J. Milton, "The use of the empirical line method to calibrate remotely sensed data to reflectance," *Int. J. Remote Sensing*, vol. 20, no. 13, pp. 2653–2662, 1999.

[143] D. A. Roberts, Y. Yamaguchi, and R. Lyon, "Comparison of various techniques for calibration of AIS data," in *Proc. 2nd Airborne Imaging Spectrometer Data Analysis Workshop*, 1986, vol. 86-35, pp. 21–30.

[144] J. E. Conel, R. O. Green, G. Vane, C. J. Brugge, and R. E. Alley, "AIS-2 radiometry and comparison of methods for the recovery of ground reflectance," in *Proc. 3rd Airborne Imaging Spectrometer Data Analysis Workshop*, 1987, vol. 95-1, pp. 49–61.

[145] B.-C. Gao, K. B. Heidebrecht, and A. F. H. Goetz, "Derivation of scaled surface reflectances from AVIRIS data," *Remote Sensing of Environment*, vol. 44, no. 2-3, pp. 165–178, 1993.

[146] F. A. Kruse, "Comparison of ATREM, ACORN, and FLAASH atmospheric corrections using low-altitude AVIRIS data of boul-

der, CO," in *Proc. 13th JPL Airborne Geoscience Workshop*, 2004, vol. 1.

[147] R. Richter and D. Schlapfer, "Geo-atmospheric processing of airborne imaging spectrometry data. Part 2: Atmospheric/topographic correction," *Int. J. Remote Sensing*, vol. 23, no. 13, pp. 2631–2649, 2002.

[148] R. N. Clark, G. A. Swayze, K. B. Heidebrecht, R. O. Green, and A. F. H. Goetz, "Calibration of surface reflectance terrestrial imaging spectrometry data: comparison of methods," in *Summaries of the 5th Annual JPL Airborne Earth Science Workshop*, 1995, vol. 95-1, pp. 41–42.

[149] M. T. Eismann, "Strategies for hyperspectral target detection in complex background environments," in *Proc. IEEE Aerospace Conference*, 2006.

[150] B. Bartlett and J.R. Schott, "Atmospheric compensation in the presence of clouds: An adaptive empirical line method (AELM) approach," *J. Applied Remote Sensing*, vol. 3, no. 1, pp. 033507, 2009.

[151] M. T. Eismann, A. D. Stocker, and N. M. Nasrabadi, "Automated hyperspectral cueing for civilian search and rescue," *Proc. IEEE*, vol. 97, no. 6, pp. 1031–1055, 2009.

[152] E. J. Ientilucci, "Spectral target detection considerations from a physical modeling perspective," in *Proc. IEEE Int. Geoscience and Remote Sensing Symposium (IGARSS)*, 2017, pp. 1320–1323.

[153] D. Tanrè, C. Deroo, P. Duhaut, M. Herman, J. J. Morcrette, J. Perbos, and P. Y. Deschamps, "Description of a computer code to simulate the satellite signal in the solar spectrum: The 5S code," *Int. J. Remote Sensing*, vol. 11, no. 4, pp. 659–668, 1990.

[154] S. M. Adler-Golden, M. W. Matthew, L. S. Bernstein, R. Y. Levine, A. Berk, S. C. Richtsmeier, P. K. Acharya, G. P. Anderson, G. Felde, J. Gardner, M. Hoke, L. S. Jeong, B. Pukall, A. Ratkowski, and H.-H. Burke, "Atmospheric correction for short-wave spectral imagery based on MODTRAN4," *Proc. SPIE*, vol. 3753, pp. 61–69, 1999.

[155] T. Haavardsholm, A. Kavara, I. Kåsen, and T. Skauli, "Improving anomaly detection with multinormal mixture models in shadow," in *Proc. IEEE Int. Geoscience and Remote Sensing Symposium (IGARSS)*, 2012, pp. 5478–5481.

[156] M. Wang and W. Shi, "Cloud masking for ocean color data processing in the coastal regions," *IEEE Trans. Geoscience and Remote Sensing*, vol. 44, no. 11, pp. 3196–3205, 2006.

[157] A. J. Brown, "Spectral curve fitting for automatic hyperspectral data analysis," *IEEE Trans. Geoscience and Remote Sensing*, vol. 44, pp. 1601–1608, 2006.

[158] A. Zare, C. Jiao, and T. Glenn, "Discriminative multiple instance hyperspectral target characterization," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 40, pp. 2342–2354, 2017.

Author Biographies

James Theiler received SB degrees in mathematics and in physics from MIT in 1981, and a PhD in physics from Caltech in 1987. He subsequently held appointments at UCSD, MIT Lincoln Laboratory, Los Alamos National Laboratory, and the Santa Fe Institute. He joined the technical staff at Los Alamos in 1994 and was named a Laboratory Fellow in 2005. He is currently an associate editor for *IEEE Transactions on Computational Imaging*. His research interests include statistical modeling, machine learning, image processing, and remote sensing.

Amanda Ziemann received BS and MS degrees in applied mathematics from Rochester Institute of Technology (RIT) in 2010 and 2011, respectively, as well as a PhD in imaging science from RIT in 2015. She was an Agnew National Security Postdoctoral Fellow at Los Alamos National Laboratory (LANL), and is currently a staff scientist in the Space Data Science and Systems Group at LANL. She is a referee for several international journals. Her research interests include remote sensing, spectral imaging, signal detection, and data fusion.

Stefania Matteoli received her BS and MS (*summa cum laude*) degrees in Telecommunications Engineering and the PhD in "Remote Sensing" from University of Pisa, Italy, in 2003, 2006, and 2010 respectively. She is currently a permanent researcher at the National Research Council of Italy within the Institute of Electronics, Computers and Telecommunication Engineering. She is Associate Editor of *IEEE Geoscience and Remote Sensing Letters*, *Journal of Applied Remote Sensing*, and *American Journal of Agricultural and Biological Sciences*. She is a referee for several international journals. In 2014 she was a recipient of the *SPIE Remote Sensing Europe Best Student Paper Award* within the "Remote Sensing of the Ocean, Sea Ice, Coastal Waters, and Large Water Regions" conference. Her research interests include signal and image processing, hyperspectral remote sensing, underwater lidar remote sensing, and atmosphere/seawater radiation transfer modeling. She is a Senior Member of the IEEE.

Marco. Diani received the Laurea degree (*cum laude*) in electronic engineering from the University of Pisa, Italy, in 1988. He is currently a Full Professor with the Italian Naval Academy, Livorno, Italy. Previously, he was an Associate Professor with the Department of Information Engineering, University of Pisa, where taught "Statistical Signal Theory" and "Design and Simulation of Remote Sensing Systems." His works covered different topics such as data fusion, signal processing in imaging radars, image classification and segmentation, object detection and tracking in infrared image sequences, and target detection and recognition in multi-hyperspectral images. His current research interests include image and signal processing with application to remote sensing. Prof. Diani is Associate Editor of the *Journal of Applied Remote Sensing*. He is a referee for several international journals. He is a member of IEEE and of SPIE.