

Subspace Based Temperature and Emissivity Separation Algorithms in LWIR hyperspectral data

N. Acito⁽¹⁾, M. Diani⁽¹⁾, G. Corsini⁽²⁾

⁽¹⁾Accademia Navale – Dip. Armi Navali – Viale Italia 72 – 57127 Livorno (Italy)

⁽²⁾Department of Information Engineering, University of Pisa – Via Caruso 16 – 56122 Pisa (Italy)

e-mail: n.acito@iet.unipi.it, m.diani@iet.unipi.it, g.corsini@iet.unipi.it

Abstract—

Index Terms— Temperature emissivity separation, Hyperspectral thermal images.

I. INTRODUCTION

Hyperspectral Remote Sensing in the Long Wave InfraRed (LWIR), provides useful information for geophysical and surveillance purposes. In general, the spectral radiance measured in the 8 – 12 μm region of the electromagnetic spectrum (LWIR) provides information about both the temperature and the chemical composition of the observed material through its emissivity. These features promise benefits in several remote sensing applications including detection of specific materials and mapping areas affected by pollutants/gases ([1], [2]).

The radiance measured by an LWIR hyperspectral sensor, is determined by the radiance emitted by the atmosphere and by the radiance leaving the observed surface (ground-leaving radiance) modified by the atmospheric path transmission. The ground-leaving radiance includes the surface emission, which depends on the temperature and emissivity of the observed material, and the atmospheric thermal emitted radiance reflected by the surface itself. Hyperspectral LWIR data exploitation requires the solution of two main problems: 1) compensation of the atmospheric effects for ground leaving radiance retrieval (Atmospheric Compensation-AC, [3]), 2) separation of the contribution due to the temperature and the emissivity in the ground leaving radiance (Temperature and Emissivity Separation-TES, [5]). These two problems are generally treated separately.

In this paper, we focus on the problem of separating the temperature from the spectral emissivity assuming that a given atmospheric compensation algorithm ([3], [4]) has provided the radiometric quantities needed to derive the ground-leaving radiance from the at-sensor radiance. TES is a challenging task because it is an ill-posed problem where the number of parameters to be estimated is higher than the number of available observations. Specifically, denoting as N_B the number of the sensor spectral channels, TES consists in estimating N_B

+1 unknowns (i.e, the N_B samples of the surface emissivity and the surface temperature) by exploiting the N_B observations corresponding to the samples of the ground leaving radiance in the sensor channels.

During the last decades, many scientists have approached this underdetermined problem from different perspectives and several algorithms were proposed to solve the problem in the case of multispectral data and, more recently, in the case of hyperspectral data. One of the most popular algorithm is that described in [5] whose improved version has been recently proposed in [6]. In [7], starting from the high spectral resolution of hyperspectral data, the analysis of the relationship between surface self-emission and atmospheric downward emitted spectral radiance in a narrow spectral region is exploited to derive the stepwise refining TES (SRTES) algorithm. SRTES uses the residue of atmospheric downward emitted spectral radiance in the calculated surface emission as a criterion, and adopts a stepwise refining method to determine both the emissivity at the location of an atmospheric emission line in a narrow spectral region and the surface temperature. In [8] the separation between temperature and emissivity is obtained by applying an iterative search algorithm that exploits a linear constraint on the spectral emissivity (Linearly Spectral Emissivity Constraint LSEC). All the cited algorithms stem from the common assumption that the emissivity of natural and man-made materials is spectrally smooth (spectral smoothness assumption).

In this paper, starting from the spectral smoothness assumption, we model the spectral emissivity as a vector lying in a subspace (emissivity subspace) having dimension K lower than the number of sensor bands (N_B). This model, in conjunction with the statistical model of the noise affecting hyperspectral data, is exploited to reformulate the TES problem as a statistical estimation framework. Approaching such a problem by means of the maximum likelihood estimation theory ([10]), we derive (Section II) a class of TES algorithms, based on the orthogonal subspace projection, which will be referred to as Subspace Based TES (SBTES), hereinafter. Several algorithms originate from the proposed approach, depending on the specific basis matrix adopted to span the emissivity subspace. In this paper, following the SBTES approach, we derive (Section III) two algorithms by exploiting two different assumptions about the emissivity subspace. Furthermore, we investigate the influence

of noise and of the emissivity subspace approximation error on the performance of the algorithms falling in the SBTES class (Section IV). The analysis is carried out by exploiting both theoretical performance bounds and simulation results. Finally, an extensive experimental analysis is proposed in order to discuss the performance of the two specific SBTES algorithms derived in this work.

II. SUBSPACE BASED TES ALGORITHMS

Let us denote as \mathbf{L} the $N_B \times 1$ vector whose components are the spectral samples of the at-sensor radiance taken at the central wavelengths $\lambda_1, \lambda_2, \dots, \lambda_{N_B}$ of the adopted sensor. In the LWIR spectral range the noisy version of the at-sensor radiance (\mathbf{L}) can be written as ([9]):

$$\begin{aligned} \mathbf{L}^{(N)} &= \mathbf{L} + \mathbf{N} \\ \mathbf{L} &= \mathbf{D}_\tau \cdot \mathbf{L}_{gl}(t, \varepsilon) + \mathbf{L}^\uparrow \end{aligned} \quad (1)$$

where \mathbf{L}^\uparrow is the $N_B \times 1$ vector of the spectral upwelling atmospheric radiance, \mathbf{D}_τ is the $N_B \times N_B$ diagonal matrix whose diagonal elements are the spectral samples of the atmospheric spectral transmittance τ , \mathbf{N} is the random noise that is modelled as a multivariate Gaussian random vector with zero mean and covariance matrix Γ_N . In (1) we denoted as $\mathbf{L}_{gl}(t, \varepsilon)$ the $N_B \times 1$ vector of the spectral ground-leaving radiance making explicit the dependence of such a quantity on the surface temperature t and the surface spectral emissivity ε . Specifically:

$$\begin{aligned} \mathbf{L}_{gl}(t, \varepsilon) &= \mathbf{D}_{BB}(t) \cdot \varepsilon + \mathbf{D}_{L^\downarrow} \cdot (1 - \varepsilon) \\ &= [\mathbf{D}_{BB}(t) - \mathbf{D}_{L^\downarrow}] \cdot \varepsilon + \mathbf{L}^\downarrow \end{aligned} \quad (2)$$

where, $\mathbf{D}_{BB}(t)$ and $\mathbf{D}_{L^\downarrow}$ are $N_B \times N_B$ diagonal matrices whose diagonal entries are the spectral radiance of a black body at temperature t (Planck's law) and the spectral downwelling atmospheric radiance (\mathbf{L}^\downarrow), respectively.

Assuming that the radiometric quantities τ , \mathbf{L}^\uparrow and \mathbf{L}^\downarrow are known and by inverting eqs. (1) and (2), the noisy vector $\mathbf{Y}(t, \varepsilon)$ is obtained as:

$$\begin{aligned} \mathbf{Y}(t, \varepsilon) &= \mathbf{D}_\tau^{-1} \cdot (\mathbf{L}^{(N)} - \mathbf{L}^\uparrow) - \mathbf{L}^\downarrow = \bar{\mathbf{D}}(t) \cdot \varepsilon + \mathbf{D}_\tau^{-1} \cdot \mathbf{N} \\ \bar{\mathbf{D}}(t) &= \mathbf{D}_{BB}(t) - \mathbf{D}_{L^\downarrow} \end{aligned} \quad (3)$$

Eq. (3) mathematically states the TES problem where we need to estimate $N_B + 1$ unknowns, i.e. t and the N_B elements of ε , by exploiting the N_B noisy observations in $\mathbf{Y}(t, \varepsilon)$. The problem is obviously ill-posed, and to come with a solution some assumptions or constraints have to be considered in order to reduce the number of unknowns or increase the number of equations.

In this work, we approach this undetermined problem, by exploiting the commonly used smoothness assumption for the emissivity spectrum. Specifically, the low pass behavior

experienced in the spectral emissivity of most of the existing materials and the consequent high degree of spectral correlation among the elements of ε , suggest that the emissivity spectrum can be considered as lying into a subspace S having rank K lower than that of the original data space (N_B). Denoting as \mathbf{U} the $N_B \times K$ matrix containing the basis vectors for the emissivity subspace S , the generic emissivity spectrum $\varepsilon \in S$ is modelled as:

$$\varepsilon = \mathbf{U} \cdot \mathbf{a} \quad (4)$$

where \mathbf{a} is the $K \times 1$ vector of the components of ε on the basis \mathbf{U} . Assuming known the basis matrix \mathbf{U} (such hypothesis will be discussed in Section III), the emissivity model in eq. (4) makes the inversion problem in (3) well-posed: the number of unknowns is reduced to $K + 1$ (the K elements of \mathbf{a} and the temperature) that is strictly less than the number N_B of noisy observations.

Notice that, according to the assumptions made, the vector $\mathbf{Y}(t, \varepsilon)$ has a multivariate Gaussian distribution with mean vector $\bar{\mathbf{D}}(t) \cdot \mathbf{U} \cdot \mathbf{a}$ and covariance matrix $\tilde{\Gamma}_N = \mathbf{D}_\tau^{-1} \cdot \Gamma_N \cdot \mathbf{D}_\tau^{-1}$. Thus, the TES problem can be approached by resorting to the statistical estimation theory. Specifically, the Maximum Likelihood (ML, [10]) estimate \hat{t} of t is obtained as (see Appendix I):

$$\begin{aligned} \hat{t} &= \arg \min_t \Psi(t) \\ \Psi(t) &= \|\mathbf{P}^\perp(t) \cdot \tilde{\mathbf{Y}}\|^2 \end{aligned} \quad (5)$$

where $\|\cdot\|$ denotes the Euclidean norm,

$$\tilde{\mathbf{Y}} = \Gamma_N^{-1/2} \cdot \mathbf{D}_\tau \cdot \mathbf{Y} \quad (6)$$

and $\mathbf{P}^\perp(t)$ is the projection matrix on the orthogonal complement of the subspace spanned by the basis matrix $\tilde{\mathbf{U}}(t)$ defined as:

$$\tilde{\mathbf{U}}(t) = \Gamma_N^{-1/2} \cdot \mathbf{D}_\tau \cdot \bar{\mathbf{D}}(t) \cdot \mathbf{U} \quad (7)$$

The estimate $\hat{\varepsilon}$ of the emissivity spectrum is obtained according to the subspace model in eq. (4) as:

$$\begin{aligned} \hat{\varepsilon} &= \mathbf{U} \cdot \hat{\mathbf{A}} \\ \hat{\mathbf{A}} &= \tilde{\mathbf{U}}(\hat{t})^+ \cdot \tilde{\mathbf{Y}} \end{aligned} \quad (8)$$

where “ $^+$ ” denotes the Moore–Penrose pseudo inverse and $\hat{\mathbf{A}}$ is the ML estimate of the vector \mathbf{a} (see Appendix I).

Eqs. (5)-(8) define a class of TES algorithms based on the assumption that ε lies onto the emissivity subspace S . Such a class of algorithms will be referred to as Subspace Based TES (SBTES), hereinafter. In practice the basis matrix \mathbf{U} of S is not known and it has to be estimated starting from certain hypotheses. Depending on the adopted basis matrix estimate, different SBTES algorithms can be obtained.

It worth noting that, in deriving the SBTES algorithms we have implicitly assumed known the noise covariance matrix Γ_N . In most practical cases, this information is not available. Several approaches can be followed to deal with such an issue. The simplest is to suppose the noise components in each band identically distributed and statistically independent so that $\Gamma_N = \sigma^2 \mathbf{I}$ (\mathbf{I} is $N_B \times N_B$ identity matrix and σ^2 is the variance of each component). More sophisticated approaches can be derived by replacing Γ_N with an estimate obtained from the available data. To this purpose, one of the numerous algorithms for noise estimation in hyperspectral data proposed in the literature ([11], [12], [13]) can be adopted.

III. EMISSIVITY SUBSPACE MATRIX DEFINITION

The specific form of the estimated basis matrix depends on the hypotheses made about the variability of the emissivity spectra within the data set of interest. Specifically, we can follow two different approaches to estimate \mathbf{U} :

1. an *unsupervised* approach, where the structure of the basis matrix is defined without any auxiliary information about the emissivity spectra;
2. a *knowledge based* approach, where the definition of the basis matrix is driven by the availability of specific information (e.g. a dictionary) about the emissivity spectra.

Depending on the adopted approach two sub-classes of SBTES algorithms can be defined: the Unsupervised SBTES (USBTES) and the Knowledge based SBTES (KSBTES). In the following, in order to clarify these concepts, we derive two algorithms one for each sub-class. Particularly, by exploiting the assumption that the generic emissivity spectrum can be approximated as a piecewise polynomial function of the sensor central wavelengths, we derive an USBTES algorithm named Polynomial SBTES (Pol-SBTES). Then, assuming that a dictionary of the emissivity spectra is available, we define a basis estimation strategy that leads to a KSBTES algorithm named Dictionary based SBTES (D-SBTES).

A. Polynomial SBTES (Pol-SBTES).

In order to derive the Pol-SBTES algorithm, we assume that in a given set of adjacent and contiguous sensor channels the emissivity spectrum can be approximated as a p -order polynomial function of the corresponding central wavelengths. Particularly, let us divide the set of all the sensor channels into M non-overlapping sections, and let us denote as $R_m = \{\lambda_{1,m}, \lambda_{2,m}, \dots, \lambda_{w,m}\}$ the set of the central wavelengths of the sensor channels falling in the m -th section. $w = N_B / M$ is the number of elements of the m -th section.

In each section the emissivity spectrum is modelled as:

$$\varepsilon(\lambda) = \sum_{l=0}^p a_l^{(m)} \cdot \lambda^l \quad \forall \lambda \in R_m \quad (9)$$

The relationship in eq. (9) can be rewritten in vector form as:

$$\boldsymbol{\varepsilon}_m = \boldsymbol{\Lambda}_m(p) \cdot \mathbf{a}_m \quad (10)$$

where $\boldsymbol{\varepsilon}_m = [\varepsilon(\lambda_{1,m}), \varepsilon(\lambda_{2,m}), \dots, \varepsilon(\lambda_{w,m})]^T$, $\mathbf{a}_m = [a_p^{(m)}, a_{p-1}^{(m)}, \dots, a_0^{(m)}]^T$ and $\boldsymbol{\Lambda}_m(p)$ is the $w \times (p+1)$ matrix defined as:

$$\boldsymbol{\Lambda}_m(p) = \begin{bmatrix} \lambda_{1,m}^p & \lambda_{1,m}^{p-1} & \dots & 1 \\ \lambda_{2,m}^p & \lambda_{2,m}^{p-1} & \dots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{w,m}^p & \lambda_{w,m}^{p-1} & \dots & 1 \end{bmatrix} \quad (11)$$

Accordingly, the entire spectrum $\boldsymbol{\varepsilon}$ can be approximated as:

$$\boldsymbol{\varepsilon} = \boldsymbol{\Lambda}(p) \cdot \mathbf{a} \quad (12)$$

where \mathbf{a} is the $M(p+1) \times 1$ vector obtained by appending all the vectors \mathbf{a}_m and $\boldsymbol{\Lambda}(p)$ is the $N_B \times M(p+1)$ matrix:

$$\boldsymbol{\Lambda}(p) = \begin{bmatrix} \boldsymbol{\Lambda}_1(p) & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Lambda}_2(p) & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \boldsymbol{\Lambda}_M(p) \end{bmatrix} \quad (13)$$

Notice that, eq. (12), subject to the constraint $M(p+1) < N_B$ (i.e. $w > p+1$), can be interpreted as the subspace representation of the emissivity spectrum (eq. (4)) where the basis matrix for the emissivity subspace is $\boldsymbol{\Lambda}(p)$. Thus, the Pol-SBTES algorithm is obtained by applying the procedure defined in eqs. (5)-(8) where the estimated basis matrix \mathbf{U} is defined as in eq. (13) ($\mathbf{U} = \boldsymbol{\Lambda}(p)$).

One of the most popular TES algorithm, named Linearly Spectral Emissivity Constraint (LSEC, [8]), can be viewed as a special case of the Pol-SBTES. It was derived by approximating the emissivity spectrum in M non-overlapping sections of the total set of sensor channels, as a linear function of the corresponding central wavelengths. Thus, such a method can be viewed as the Pol-SES algorithm when a first order polynomial approximation ($p=1$) is assumed. However, it is worth noting that the cost function $\Psi(t)$ adopted by the Pol-SES algorithm accounts for the noise covariance matrix Γ_N (eq. (7)), whereas the LSEC in its original formulation does not account for random noise affecting the at-sensor radiance. The Pol-SES algorithm with $p=1$ coincides with the LSEC when the assumption $\Gamma_N = \sigma^2 \mathbf{I}$ is made.

B. Dictionary based SBTES (D-SBTES).

When a dictionary \mathbf{V} containing the spectral emissivities of several natural or man-made materials is available, it can be exploited to estimate a basis matrix for the emissivity subspace, thus obtaining the D-SBTES algorithm. Several strategies can be defined to extract \mathbf{U} from the dictionary \mathbf{V} . In this work, we exploit the idea that in the LWIR portion of the electromagnetic spectrum, the variability of the dictionary atoms \mathbf{v}_g , $g=1, \dots, G$ around their spectral mean value μ_g can be well represented on a subspace \bar{S} having rank \bar{K} , i.e.:

$$\tilde{\mathbf{v}}_g \approx \bar{\mathbf{U}} \cdot \mathbf{b}_g \quad \forall g=1, \dots, G \quad (14)$$

where, G is the number of dictionary atoms, $\tilde{\mathbf{v}}_g$ is the g -th dictionary atom after spectral mean removal ($\tilde{\mathbf{v}}_g = \mathbf{v}_g - \mu_g$) and

$\bar{\mathbf{U}}$ is the basis matrix of $\bar{\mathcal{S}}$. $\bar{\mathbf{U}}$ may be obtained by searching for the basis matrix that solves the following constrained minimum mean square error problem:

$$\begin{aligned} \bar{\mathbf{U}} = \arg \min & \left\{ \frac{1}{G} \sum_{g=1}^G \|\bar{\mathbf{v}}_g - \bar{\mathbf{U}} \cdot \mathbf{b}_g\|^2 \right\} \\ \text{s.t.} : & \frac{1}{G} \sum_{g=1}^G \|\bar{\mathbf{v}}_g - \bar{\mathbf{U}} \cdot \mathbf{b}_g\|^2 < \eta \cdot \frac{1}{G} \sum_{g=1}^G \|\bar{\mathbf{v}}_g\|^2 \quad \text{with } 0 < \eta < 1 \end{aligned} \quad (15)$$

The solution of this problem is obtained by using the well-known Singular Value Decomposition (SVD, [14]) algorithm, i.e. retaining the singular vectors that address a fraction of the mean atoms power ($\frac{1}{G} \sum_{g=1}^G \|\bar{\mathbf{v}}_g\|^2$) greater than $1 - \eta$. To have a good subspace representation, η must be close to zero (for instance $\eta = 0.01$)

It is worth noting that, to derive the basis matrix \mathbf{U} accounting for the mean spectrum of each atom \mathbf{v}_g (and of the generic spectral emissivity $\boldsymbol{\varepsilon}$), the $N_B \times 1$ vector with all unit entries ($\mathbf{1}$) has to be added to the basis matrix $\bar{\mathbf{U}}$, thus obtaining $\mathbf{U} = [\bar{\mathbf{U}}, \mathbf{1}]$.

The dictionary \mathbf{V} can be constructed by using in-field emissivity measurements or by exploiting spectral libraries made available for the scientific community in the last years. Examples of open access spectral libraries are: the ASTER spectral library ([15]) hosted by Jet Propulsion Laboratory (JPL), the Moderate Resolution Imaging Spectrometer (MODIS) UCSB Emissivity Library made available by the Institute for Computational Earth System Science of the University of California ([16]), and the London Urban Micromet data Archive (LUMA, [17]) including emissivity spectra of urban materials.

IV. PERFORMANCE ANALYSIS

This Section has two main objectives: the discussion of the performance of a generic SBTES algorithm and the analysis of the specific estimation performance attained by two of the SBTES algorithms introduced in Section III, i.e. the Pol-SBTES algorithm with $p=1$ (Pol-SBTES-1) and the D-SBTES algorithm. The general discussion about the temperature and the emissivity estimators is focused on the impacts of noise and of the subspace approximation error and is carried out by exploiting both theoretical results and experimental evidences. Analytical expressions (derived in Appendix II) are adopted to obtain bounds on the accuracy of both the temperature and the emissivity estimators. Performance analysis of the Pol-SBTES-1 the D-SBTES algorithms, is carried out on simulated data.

Since simulated data play an important role throughout this Section, we start by describing the adopted simulation strategy in Section IV.A. Then, in Section IV.B, we discuss the performance of a generic SBTES algorithm. Finally, in Section

IV.C, we present and analyze the results obtained on simulated data by Pol-SBTES-1 and D-SBTES.

A. Simulation strategy.

At-sensor radiance data are simulated according to the simulation strategy outlined in the block diagram of Figure 1.

Starting from the dictionary \mathbf{V}^{TOT} obtained by merging the ASTER, the MODIS and the LUMA spectral libraries, an emissivity spectrum $\boldsymbol{\varepsilon}$ is extracted. According to the radiative transfer model in eqs. (1) and (2), the radiometric quantities τ , L^\uparrow and L^\downarrow are generated by MODTRAN for a given set of the geometric and atmospheric parameters characterizing the acquisition environment. We refer to airborne applications and consider a flight altitude of 1 km.

The radiometric quantities, the surface temperature t and the emissivity spectrum $\boldsymbol{\varepsilon}$ are combined according to eqs. (1)-(2) to obtain the at-sensor spectral radiance L^* . This spectrum is properly transformed in the sensor spectral domain (\mathbf{L}) accounting for the sensor specifications. For this purpose, we refer to the Hyperspectral Thermal Emission Spectrometer (HyTES, [21]) that is an airborne imaging spectrometer with 256 spectral channels between 7.5 and 12 μm , with a Full Width Half Maximum (FWHM) of 35 nm and a spectral sampling of 17 nm. Generally, the bands covering the spectral range from 7.5 to 8 μm affected by strong atmospheric absorption are discarded. The resulting number of spectral bands is $N_B = 229$.

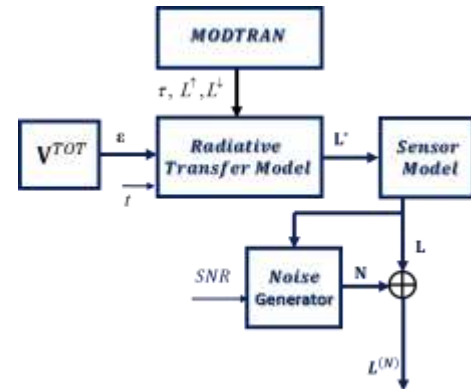


Figure 1. Block diagram of the simulation strategy.

Finally, the noisy at-sensor radiance $L^{(N)}$ is obtained by considering a Gaussian additive noise with zero mean and diagonal covariance matrix. The method adopted to generate the noise covariance matrix attempts to obtain realistic noise realizations and deserves a detailed description. Specifically, we assume that the instrument is photon noise limited so as to have a noise variance per band proportional to the at-sensor radiance ([21]):

$$\begin{aligned} \sigma_N^2(\lambda_l, L) &= \psi_{ph}(\lambda_l) \cdot L(\lambda_l) \\ \psi_{ph}(\lambda_l) &= \frac{\sigma^2}{\lambda_l} \end{aligned} \quad (16)$$

where l is the band index, λ_l is the central wavelength of the l -th band, $\sigma_N^2(\lambda_l, L)$ is the noise variance in the l -th band, and σ^2 is a wavelength independent multiplicative factor (the quantum efficiency is assumed constant). Accordingly, the noise covariance matrix can be expressed as $\Gamma_N = \sigma^2 \cdot \Gamma_0$, Γ_0 being the diagonal matrix with l -th diagonal entry $\frac{L(\lambda_l)}{\lambda_l}$. The Signal to Noise Ratio (SNR) defined as:

$$\begin{aligned} SNR(L) &= \frac{1}{N_B} \left\| \Gamma_N^{-1/2} \cdot \mathbf{L} \right\|^2 = \frac{SNR_0(L)}{\sigma^2} \\ SNR_0(L) &= \frac{1}{N_B} \left\| \Gamma_0^{-1/2} \cdot \mathbf{L} \right\|^2 \end{aligned} \quad (17)$$

can be set by choosing the multiplicative factor σ^2 .

B. On the performance of SBTES algorithms.

Assuming known the radiometric quantities τ , \mathbf{L}^\dagger and \mathbf{L}^\ddagger , the performance of a generic SBTES algorithm depends on several factors such as: the ability of the adopted basis matrix \mathbf{U} to properly address the emissivity spectrum ϵ ; the random noise affecting the at-sensor radiance data; the specific values of the surface temperature and of the emissivity spectrum.

In the following we discuss the impact of some of these factors on the performance of both the estimators \hat{T} and $\hat{\epsilon}$.

Let us start with \hat{T} and let us denote as $\Delta\epsilon$ the error due to the approximation of ϵ on the subspace spanned by \mathbf{U} , i.e. :

$$\epsilon = \mathbf{U} \cdot \mathbf{a} + \Delta\epsilon \quad (18)$$

Letting $\tilde{\mathbf{D}}(t) = \Gamma_N^{-1/2} \cdot \mathbf{D}_r \cdot \tilde{\mathbf{D}}(t)$, the random vector $\tilde{\mathbf{Y}}$ in eq. (6) can be rewritten as:

$$\begin{aligned} \tilde{\mathbf{Y}} &= \tilde{\mathbf{U}}(t_r) \cdot \mathbf{a} + \tilde{\mathbf{D}}(t_r) \cdot \Delta\epsilon + \Gamma_N^{-1/2} \cdot \mathbf{N} = \\ &= \tilde{\mathbf{U}}(t_r) \cdot \mathbf{a} + \tilde{\mathbf{D}}(t_r) \cdot \Delta\epsilon + \mathbf{W} \end{aligned} \quad (19)$$

where t_r is the true temperature of the observed surface, and the noise term $\mathbf{W} = \Gamma_N^{-1/2} \cdot \mathbf{N}$ is a zero mean Gaussian distributed vector with identity covariance matrix.

The cost function $\Psi(t)$ in eq. (5) can be rewritten so as to highlight the contribution of the subspace approximation error ($\Omega_{\Delta\epsilon}(t, t_r)$) and that of the random noise ($\Omega_{\mathbf{W}}(t)$):

$$\begin{aligned} \Psi(t) &= \left\| \mathbf{P}^\perp(t) \cdot \tilde{\mathbf{Y}} \right\|^2 = \left\| \Omega(t, t_r) + \Omega_{\Delta\epsilon}(t, t_r) + \Omega_{\mathbf{W}}(t) \right\|^2 \\ \Omega(t, t_r) &= \mathbf{P}^\perp(t) \cdot \tilde{\mathbf{U}}(t_r) \cdot \mathbf{a} \\ \Omega_{\Delta\epsilon}(t, t_r) &= \mathbf{P}^\perp(t) \cdot \tilde{\mathbf{D}}(t_r) \cdot \Delta\epsilon \\ \Omega_{\mathbf{W}}(t) &= \mathbf{P}^\perp(t) \cdot \mathbf{W} \end{aligned} \quad (20)$$

In the following, we will analyze the effects of the approximation error and the noise separately. Notice that, in absence of both approximation error ($\Delta\epsilon = \mathbf{0}$ and then $\Omega_{\Delta\epsilon}(t) = \mathbf{0}$) and noise ($\Omega_{\mathbf{W}}(t) = \mathbf{0}$), $\Psi(t) = \left\| \Omega(t, t_r) \right\|^2$ is a convex function of t having its minimum in $t = t_r$. It has always positive values except for $t = t_r$ where it is zero, i.e. $\mathbf{P}^\perp(t = t_r)$ projects the data $\tilde{\mathbf{U}}(t_r) \cdot \mathbf{a}$ on the orthogonal complement of the subspace spanned

by $\tilde{\mathbf{U}}(t_r)$. In such a case the generic SBTES algorithm, provides the correct value of the surface temperature. When only the approximation error is present $\Psi(t)$ can be expressed as:

$$\begin{aligned} \Psi(t) &= \left\| \Omega(t, t_r) \right\|^2 + \Phi(t, \Delta\epsilon) \\ \Phi(t, \Delta\epsilon) &= 2 \cdot \Omega(t, t_r)^T \cdot \Omega_{\Delta\epsilon}(t, t_r) + \left\| \Omega_{\Delta\epsilon}(t, t_r) \right\|^2 \end{aligned} \quad (21)$$

The scalar function $\Phi(t, \Delta\epsilon)$ depends on $\Delta\epsilon$ and has the effect of deviating the minimum of $\Psi(t)$ from $t = t_r$, thus inducing an estimation error that depends on the specific value of the vector $\Delta\epsilon$ and, of course, on the adopted basis matrix. We use an example on simulated data to clarify this. Specifically, we refer to the HyTES sensor ($N_B = 229$) and we consider the emissivity spectrum ϵ_r of the dictionary \mathbf{V}^{TOT} corresponding to the man-made material named ‘‘Oxidized Galvanized Steel Metal’’ (Figure 2 a)). According to the simulation strategy in Sub-section IV.A, we obtain the at-sensor radiance for a given atmosphere (generated by MODTRAN) and a surface temperature of 30 °C ($t_r = 30^\circ\text{C}$).

For this example, we consider the D-SBTES algorithm where the basis matrix \mathbf{U} is estimated from \mathbf{V}^{TOT} according to the SVD based procedure in Sub-Section II.B. Setting $\eta = 0.01$ we obtain \mathbf{U} having rank $K = 8$. The spectrum ϵ_r is viewed as the sum of two orthogonal components: that lying on the subspace spanned by \mathbf{U} ($\mathbf{U} \cdot \mathbf{U}^+ \cdot \epsilon_r$, Figure 2 a)) and the approximation error $\Delta\epsilon$.

According to eq. (19), in Figure 2 b) we plot $\left\| \Omega(t, t_r) \right\|^2$, $\Phi(t, \Delta\epsilon)$ and $\Psi(t)$ for t ranging from 20°C to 40°C. Note that $\left\| \Omega(t, t_r) \right\|^2$ is indeed a convex function of t and has its minimum value in $t = t_r$ (green point). On the contrary, in this example, $\Phi(t, \Delta\epsilon)$ is a decreasing monotonic function of t , and its effect on the cost function $\Psi(t)$ is to shift the value of t where the minimum is reached (red point in Figure 2 b)). In this specific case $\Psi(t)$ attains its minimum in $t = 33^\circ\text{C}$ ($\hat{T} = 33^\circ\text{C}$)

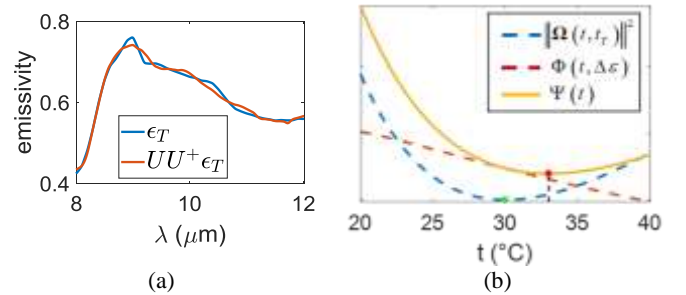


Figure 2. (a) Representation of the emissivity spectrum ϵ_r and of its approximation on the basis \mathbf{U} obtained by setting $\eta = 0.01$ (\mathbf{U}^+ denotes the pseudo inverse of \mathbf{U}); (b) representation of the three terms $\left\| \Omega(t, t_r) \right\|^2$, $\Phi(t, \Delta\epsilon)$ and $\Psi(t)$ obtained by using the basis matrices \mathbf{U} .

The error on \hat{T} decreases with decreasing values of $\Delta\epsilon$ which depends on the specific basis matrix \mathbf{U} . To improve the subspace based emissivity approximation and reduce $\Delta\epsilon$, more

basis vectors can be added to \mathbf{U} (by increasing its rank K). In Figure 4 we plot the functions $\|\Omega(t, t_r)\|^2$, $\Phi(t, \Delta\varepsilon)$ and $\Psi(t)$ corresponding to the basis matrices obtained by setting $\eta = 0.008$ ($K = 9$, Figure 4 a) and $\eta = 0.0004$ ($K = 24$, Figure 4 b)). For completeness, in Figure 3 we plot the subspace approximation of ε_T on the two basis matrices ($K = 9$, Figure 3 a) and $K = 24$, Figure 3 b)).

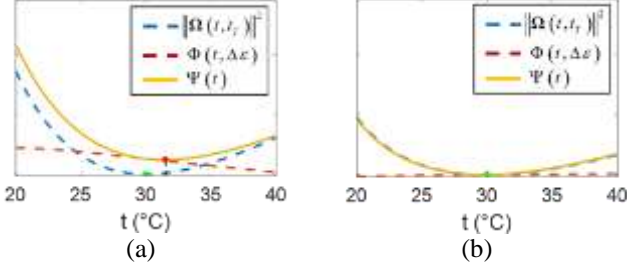


Figure 3. Representation of three terms $\|\Omega(t, t_r)\|^2$, $\Phi(t, \Delta\varepsilon)$ and $\Psi(t)$ obtained by using two different basis matrices \mathbf{U} with rank: a) $K = 9$, b) $K = 24$, respectively.

Notice that, the effects of the approximation error already decreases when the rank is increased by one. In Figure 4 a) the value of t where $\Psi(t)$ has its minimum is 31.5. The error in the temperature estimate vanishes when the basis matrix with $K = 24$ is adopted (Figure 4 b)). In this case, $\Phi(t, \Delta\varepsilon)$ is negligible and $\Psi(t)$ is actually coincident with $\|\Omega(t, t_r)\|^2$. Such a basis matrix provides a very low $\Delta\varepsilon$ as witnessed by the approximated spectrum in Figure 3 b).

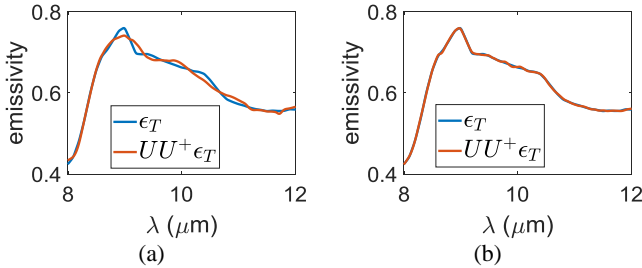


Figure 4. Representation of the emissivity spectrum ε_T and of its approximation on the basis $\hat{\mathbf{U}}$ obtained by setting: a) $\eta = 0.008$ ($K = 9$) and b) $\eta = 0.0004$ ($K = 24$).

In order to analyze the performance of the SBTES algorithms with respect to noise, let us assume $\Delta\varepsilon = \mathbf{0}$ so as to have $\Psi(t) = \|\Omega(t, t_r) + \Omega_w(t)\|^2$. The discussion about the noise impact can be performed by considering the Cramer-Rao Lower Bound (CRLB, [10]) for the estimation of the deterministic parameter t . The CRLB is the lower bound for the variance of any estimator of t . Thus, the CRLB provides insights about the maximal theoretical accuracy that can be achieved by temperature estimation methods.

In our case the analysis of the CRLB is even more interesting because \hat{t} in eq. (5) is the ML estimator of t . Thus, it has the asymptotical properties of being unbiased and achieving the CRLB ([10]).

The analytical expression of the CRLB is derived in Appendix II, whereas in this Section, we exploit such an expression to discuss the dependence of the estimation performance on the parameters that define the approached problem. The CRLB for the estimators of t based on the subspace model in eqs. (3) and (4) can be written as:

$$CRLB_t(t_r, \mathbf{a}) = \frac{1}{\|\mathbf{P}^\perp(t_r) \cdot \bar{\mathbf{D}}^{(1)}(t_r) \cdot \tilde{\varepsilon}\|^2} \quad (22)$$

$$\bar{\mathbf{D}}^{(1)}(t_r) = \Gamma_N^{-1/2} \cdot \mathbf{D}_t \cdot \mathbf{D}_{BB}^{(1)}(t_r)$$

$$\tilde{\varepsilon} = \mathbf{U} \cdot \mathbf{a}$$

where, t_r is the true value of the temperature, $\mathbf{D}_{BB}^{(1)}(t_r)$ is the diagonal matrix whose l -th entry $d_{BB}^{(1)}(\lambda_l, t_r)$ is the derivative of the Plank's law with respect to the temperature at the central wavelength λ_l , i.e.:

$$d_{BB}^{(1)}(\lambda_l, t_r) = \left. \frac{\partial}{\partial t} d_{BB}(\lambda_l, t) \right|_{t=t_r} \quad (23)$$

and $\tilde{\varepsilon}$ is the component of ε lying on the subspace spanned by \mathbf{U} .

The expression of the CRLB in eq. (22) shows that the temperature estimation accuracy depends on the noise covariance matrix Γ_N (and then on the SNR) and on the specific value of t_r . To make explicit the dependence on the SNR let us consider the assumption about the noise covariance matrix in Section IV.A: $\Gamma_N = \sigma^2 \cdot \Gamma_0$. Letting:

$$\tilde{\mathbf{U}}_0(t) = \Gamma_0^{-1/2} \cdot \mathbf{D}_t \cdot \bar{\mathbf{D}}^{(1)}(t) \cdot \mathbf{U} \quad (24)$$

$$\bar{\mathbf{D}}_0^{(1)}(t) = \Gamma_0^{-1/2} \cdot \mathbf{D}_t \cdot \mathbf{D}_{BB}^{(1)}(t)$$

and then:

$$\tilde{\mathbf{U}}(t) = \frac{1}{\sigma} \cdot \tilde{\mathbf{U}}_0(t) \quad (25)$$

$$\bar{\mathbf{D}}^{(1)}(t) = \frac{1}{\sigma} \cdot \bar{\mathbf{D}}_0^{(1)}(t)$$

it can be easily proved that $\mathbf{P}^\perp(t_r)$ does not depend on σ , and

$$CRLB_t(t_r, \mathbf{a}) = \sigma^2 \cdot CRLB_{t,0}(t_r, \mathbf{a})$$

$$CRLB_{t,0}(t_r, \mathbf{a}) = \frac{1}{\|\mathbf{P}^\perp(t_r) \cdot \bar{\mathbf{D}}_0^{(1)}(t_r) \cdot \tilde{\varepsilon}\|^2} \quad (26)$$

Since σ^2 is inversely proportional to SNR (eq. (17)), it is evident from eq. (26) that $CRLB_t(t_r, \mathbf{a})$ increases by decreasing SNR .

To numerically analyze such a dependence, we use an example similar to that introduced in the previous part of this section. Namely, we consider the D-SBTES algorithm and the spectrum ε_T , the at-sensor radiance is obtained by specifying the atmosphere and the surface temperature t_r , and the coefficient σ^2 is chosen so as to obtain the desired SNR . According with the assumption $\Delta\varepsilon = \mathbf{0}$, the vector \mathbf{a} is set to $\mathbf{U}^+ \cdot \varepsilon_T$.

We derive the numerical values of $CRLB_t(t_r, \mathbf{a})$ for various SNR and t_r , when the basis matrix \mathbf{U} is obtained by setting $\eta = 0.01$ ($K = 8$). In Figure 5 a) the square root values of $CRLB_t(t_r, \mathbf{a})$ for

SNR ranging from 20 to 60 dB are shown. We can note that, as expected, the accuracy improves for increasing SNR : the lower bound for the standard deviation of the temperature estimator decreases as SNR increases for all the considered surface temperatures. We can also note that, we obtain better performance with lower values of the surface temperature.

The expression in eq. (20) shows that the impact of noise on the estimation performance depends also on the rank K of the emissivity subspace basis matrix. Specifically, it shows that $CRLB_i(t_r, \mathbf{a})$ is inversely proportional to the Euclidean norm of the projection of $\bar{\mathbf{D}}^{(i)}(t_r) \cdot \bar{\boldsymbol{\varepsilon}}$ onto the subspace spanned $\mathbf{P}^\perp(t_r)$ that has rank $N_B - K$. When $N_B - K$ decreases (K increases), such a norm decreases too. As a consequence, it is expected that $CRLB_i(t_r, \mathbf{a})$ increases with K . This is evident in Figure 5 b) that plots the values of the square root of $CRLB_i(t_r, \mathbf{a})$ for different values of K and SNR for the previously introduced example.

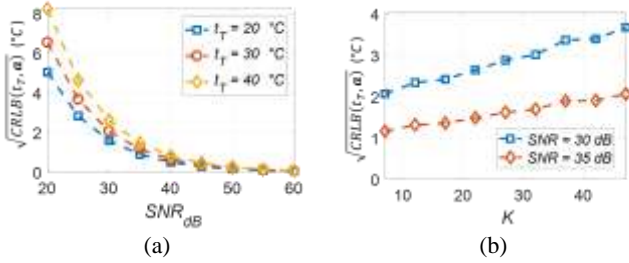


Figure 5. (a) Square root of the CRLB for the surface temperature estimator obtained by using the D-SBTES algorithm with a basis matrix having rank $K = 8$. (b) Square root of the CRLB obtained by varying the rank of emissivity subspace basis matrix.

The fact that the impact of noise on the performance of the SBTES algorithms augments as K increases is not quite intuitive. In fact, by augmenting K the statistical power of the noise term $\boldsymbol{\Omega}_w(t)$ in eq. (20), decreases because it is obtained by projecting the white noise vector \mathbf{W} on the subspace with rank $N_B - K$. So, we would expect an improvement in the estimation performance as K increases. Actually, by increasing K , the curvature of $\|\boldsymbol{\Omega}(t, t_r)\|^2$ decreases (see Figure 6) and its minimum becomes less distinguishable in the presence of noise even if the value of the noise variance decreases.

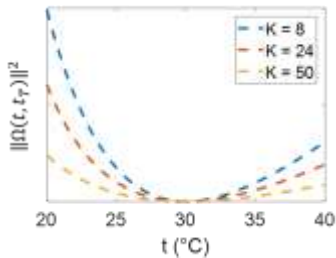


Figure 6. $\|\boldsymbol{\Omega}(t, t_r)\|^2$ for $t_r = 30^{\circ}\text{C}$ obtained by basis matrices with different ranks: $K = 8$, $K = 24$ and $K = 50$.

Finally, the expression of the CRLB in eq. (20), leads to conclude that SBTES algorithms provide in general the worst performance on low-emissivity materials. In fact, $CRLB(t_r)$ is inversely proportional to $\|\mathbf{P}^\perp(t_r) \cdot \bar{\mathbf{D}}^{(i)}(t_r) \cdot \bar{\boldsymbol{\varepsilon}}\|^2$, which is lower

than $\|\bar{\boldsymbol{\varepsilon}}\|^2$. So, at low values of $\|\bar{\boldsymbol{\varepsilon}}\|^2$ correspond high values of $CRLB_i(t_r, \mathbf{a})$. This is exemplified in Figure 7 a), which shows the values of $\sqrt{CRLB_i(t_r, \mathbf{a})}$ for three materials of the \mathbf{V}^{TOT} dictionary having very low emissivity (ε_1), low emissivity (ε_2) and high emissivity (ε_3), respectively. These numerical results have been obtained by applying the D-SBTES algorithm with $\eta = 0.01$ ($K = 8$) and by setting $t_r = 30^{\circ}\text{C}$ and $SNR = 30$ dB. For completeness, in Figure 7 b) we plot the three spectra. Figure 7 a) shows that, performance actually gets worse when the emissivity decreases.

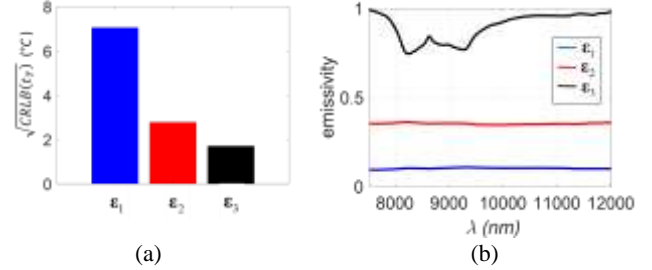


Figure 7. (a) Square root of the CRLB for the surface temperature estimator obtained by applying the D-SBTES algorithm to three spectra and by setting $\eta = 0.01$ ($K = 8$), $t_r = 30^{\circ}\text{C}$ and $SNR = 30$ dB; (b) emissivity spectra of the materials considered in the example.

To discuss the performance of the estimator of the spectral emissivity in eq. (8), we analyze the relative mean square error $\gamma(t, \boldsymbol{\varepsilon})$ defined as:

$$\gamma(t, \boldsymbol{\varepsilon}) = \frac{E\{\|\hat{\boldsymbol{\varepsilon}} - \boldsymbol{\varepsilon}\|^2\}}{\|\boldsymbol{\varepsilon}\|^2} \quad (27)$$

where E is the statistical expectation operator. Assuming $\hat{\mathbf{a}}$ as an unbiased estimator of \mathbf{a} (the ML estimator is asymptotically unbiased), and considering the model in eq. (18), $\gamma(t, \boldsymbol{\varepsilon})$ is rewritten as:

$$\gamma(t, \boldsymbol{\varepsilon}) = \gamma_w(t, \mathbf{a}) + \frac{\|\Delta\boldsymbol{\varepsilon}\|^2}{\|\boldsymbol{\varepsilon}\|^2} \quad (28)$$

$$\gamma_w(t, \mathbf{a}) = \frac{E\left\{\left\|\mathbf{U} \cdot (\hat{\mathbf{A}} - \mathbf{a})\right\|^2\right\}}{\|\boldsymbol{\varepsilon}\|^2}$$

where, $\gamma_w(t, \mathbf{a})$ depends on noise and on $\Delta\boldsymbol{\varepsilon}$ through the estimate of t .

Again, we discuss separately the impact of noise and of the approximation error. In absence of noise, it is evident from eq. (28) that $\gamma(t, \boldsymbol{\varepsilon})$ increases with increasing values of $\|\Delta\boldsymbol{\varepsilon}\|$. Performance improvement is obtained by reducing the approximation error, for instance, by augmenting the rank K of the matrix \mathbf{U} .

To discuss the impact of noise when $\Delta\boldsymbol{\varepsilon} = \mathbf{0}$, we exploit the results derived in Appendix II. Specifically, we consider that for any unbiased estimator of \mathbf{a} (and $\boldsymbol{\varepsilon}$), a lower bound for $\gamma_w(t, \mathbf{a})$ is:

$$\gamma_{LB}(t_T, \mathbf{a}) = \sigma^2 \cdot \frac{\mathcal{G}_{LB,0}(t_T, \mathbf{a})}{\|\boldsymbol{\varepsilon}\|^2}$$

$$\mathcal{G}_{LB,0}(t_T, \mathbf{a}) = \text{tr}\{\mathbf{B}_0(t_T)\} + \text{CRLB}_{i,0}(t_T, \mathbf{a}) \cdot \|\mathbf{U} \cdot \tilde{\mathbf{U}}_0(t_T)^+ \cdot \mathbf{b}_0(t_T, \mathbf{a})\|^2 \quad (29)$$

with:

$$\mathbf{B}_0(t_T) = \mathbf{U} \cdot [\tilde{\mathbf{U}}_0(t_T)^T \cdot \tilde{\mathbf{U}}_0(t_T)]^{-1} \cdot \mathbf{U}^T$$

$$\mathbf{b}_0(t_T, \mathbf{a}) = \tilde{\mathbf{D}}_0^{(i)}(t_T) \cdot \mathbf{U} \cdot \mathbf{a} \quad (30)$$

Notice that, in writing the expression of $\gamma_{LB}(t_T, \mathbf{a})$ we have implicitly assumed the noise covariance matrix model presented in Section IV.A in order to make explicit the dependence of such a lower bound on SNR . The general expression of $\gamma_{LB}(t_T, \mathbf{a})$ is derived by the results in Appendix II. Here again we can say that, the analysis of $\gamma_{LB}(t_T, \mathbf{a})$ gives insights about the best theoretical accuracy of the emissivity estimator. We would like to stress that, in our case, such an analysis is more significant because of the asymptotical properties of the adopted ML estimator.

The expression of $\gamma_{LB}(t_T, \mathbf{a})$ highlights that, as expected, the performance of the emissivity estimator decreases when that of the temperature estimator worsens. In fact, $\gamma_{LB}(t_T, \mathbf{a})$ is directly proportional to $\text{CRLB}_i(t_T, \mathbf{a})$.

Furthermore, eq. (29) also shows that $\gamma_{LB}(t_T, \mathbf{a})$ increases as SNR decreases (σ^2 increases). We give a numerical evidence by considering the previously introduced example concerning the spectrum ε_T . Specifically, in Figure 8 a) we plot the values of $\gamma_{LB}(t_T, \mathbf{a})$ (in percent) obtained by applying the D-SBTES algorithm with $K=8$ and for SNR ranging from 20 to 60 dB. Such a figure shows the performance improvement achieved with increasing SNR . Figure 8 a), also shows that the performance depends on the true temperature values and that the accuracy slightly improves by decreasing t_T .

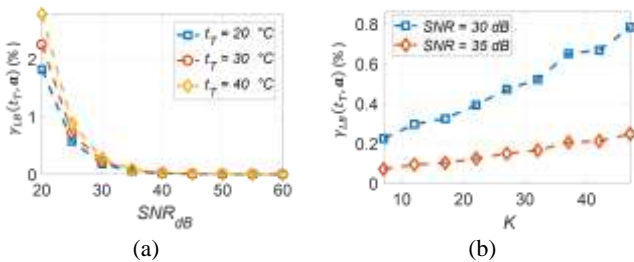


Figure 8. (a) Values of $\gamma_{LB}(t_T, \mathbf{a})$ obtained by using the D-SBTES algorithm with a basis matrix having rank $K=8$ for SNR ranging from 20 to 60 dB. (b) Values of $\gamma_{LB}(t_T, \mathbf{a})$ obtained by varying the rank of the emissivity subspace basis matrix.

As to the dependence of $\gamma_{LB}(t_T, \mathbf{a})$ on K , we observe that, the matrix $\mathbf{B}_0(t_T)$ is a positive semidefinite matrix, so by adding columns to \mathbf{U} and increasing its rank K (and that of $\tilde{\mathbf{U}}_0(t_T)$), the sum of its eigenvalues (i.e. its trace) augments. $\text{CRLB}_{i,0}(t_T, \mathbf{a})$ has been proved to increase with K . Increasing trend with K is expected also for the term $\|\mathbf{U} \cdot \tilde{\mathbf{U}}_0(t_T)^+ \cdot \mathbf{b}_0(t_T, \mathbf{a})\|^2$. In fact, it is

the squared Euclidean norm of the vector obtained by the linear combination of the columns of \mathbf{U} with coefficients given by $\tilde{\mathbf{U}}_0(t_T)^+ \cdot \mathbf{b}_0(t_T, \mathbf{a})$. Adding components to the basis \mathbf{U} (and so to $\tilde{\mathbf{U}}_0(t_T)$) we add components to the aforementioned linear combination and, consequently, its squared Euclidean norm is expected to increase. Thus, we can conclude that $\gamma_{LB}(t_T, \mathbf{a})$ increases by augmenting the rank of the basis matrix. Numerical evidence is given in Figure 8 b) where the values of $\gamma_{LB}(t_T, \mathbf{a})$ obtained by applying the D-SBTES in the case of the spectrum ε_T , are plotted for different values of K ranging from 8 to 50 and for two different values of SNR .

Finally, eq. (29) shows that the value of $\gamma_{LB}(t_T, \mathbf{a})$ strictly depends on the norm of the emissivity spectrum ($\mathbf{U} \cdot \mathbf{a}$). Specifically, we note that for low emissivity spectra, $\text{CRLB}_{i,0}(t_T, \mathbf{a})$ is expected to have high value, conversely, $\|\mathbf{U} \cdot \tilde{\mathbf{U}}_0(t_T)^+ \cdot \mathbf{b}_0(t_T, \mathbf{a})\|^2$ has low values. Thus, the specific relationships between the performance of the emissivity estimator and the norm of the emissivity spectrum is not pretty clear. To clarify this point, we derived the values of $\gamma_{LB}(t_T, \mathbf{a})$ in the case of D-SBTES algorithm (with $K=8$) for all the spectra in the \mathbf{V}^{TOT} dictionary.

In Figure 9, we plot the values of $\gamma_{LB}(t_T, \mathbf{a})$ as a function of the mean Euclidean norm of the emissivity spectra when $t_T = 30^\circ\text{C}$ and $SNR=30$ dB. Figure 9 clearly shows that $\gamma_{LB}(t_T, \mathbf{a})$ has the highest values for the lowest emissivity spectra. Thus, in general, we have the worst performance for the lowest emissivity spectra.

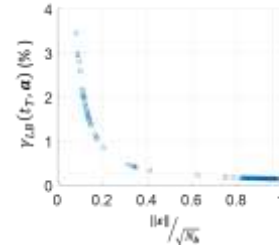


Figure 9. Values of $\gamma_{LB}(t_T, \mathbf{a})$ versus the mean Euclidean norm of the emissivities in the \mathbf{V}^{TOT} dictionary. Results obtained by applying the D-SBTES algorithm with $K=8$ and by setting $t_T = 30^\circ\text{C}$ and $SNR=30$ dB.

To conclude this section, we would like to stress that, though the effects of noise and $\Delta\varepsilon$ have been analyzed separately, actually they act jointly. The analysis has shown that the performance of SBTES algorithms strongly depends on the basis matrix \mathbf{U} adopted for addressing the emissivity subspace. Specifically, we have shown that \mathbf{U} has to provide a good representation of the emissivity spectra in the analyzed data set thus having low representation error vectors ($\Delta\varepsilon$). High rank matrices are generally suitable to meet this requirement. On the other hand, the rank of \mathbf{U} must be as small as possible in order to reduce the performance degradation due to random noise. Thus, \mathbf{U} be carefully designed so as to provide a “good” and “compact” representation of the emissivity spectra populating the data set of interest.

C. Experimental results on simulated data.

In this Sub-Section, the combined effects of noise and of the representation error are numerically analyzed on simulated data. To show that the performance of a given SBTES algorithm depends on the adopted basis matrix, we derive numerical results for two of the algorithms introduced in Section III: Pol-SBTES-1 and D-SBTES.

The test set is simulated according to the procedure summarized in Section IV.A. Specifically, for a given atmosphere and a given surface temperature t_r , we generated the at-sensor radiance for the 357 spectra ε_q ($q=1,\dots,357$) in the \mathbf{V}^{TOT} dictionary. For each radiance spectrum we derived 100 noisy at-sensor radiance samples by adding 100 realizations of noise vectors generated according to the model introduced in Section IV.A. The resulting number of samples in the test set was 35700. As concerns the atmosphere, results discussed in the following were derived by considering a flight altitude of 1 Km and the mid-latitude summer MODTRAN atmospheric model. On the basis of the root mean Euclidean norm:

$$\rho_q = \sqrt{\frac{1}{N_B} \|\varepsilon_q\|^2} \quad (31)$$

the spectra ε_q are grouped in three classes: high emissivity spectra (“High”, $\rho_q \geq 0.6$), low emissivity spectra (“Low”, $0.1 \leq \rho_q < 0.6$) and very low emissivity spectra (“Very Low” $\rho_q < 0.1$). For each class, the following performance indexes are computed:

$$\begin{aligned} \bar{\gamma}(t_r) &= \frac{1}{N_Q} \sum_Q \gamma(t_r, \varepsilon_q) \\ \bar{\sigma}_T(t_r) &= \frac{1}{N_Q} \sum_Q \sigma_T(t_r, \varepsilon_q) \\ \sigma_T(t_r, \varepsilon_q) &= \sqrt{E \left\{ \left[\hat{T}(t_r, \varepsilon_q) - t_r \right]^2 \right\}} \end{aligned} \quad (32)$$

In eq. (32), Q denotes the set of indexes addressing the spectra of a given class, N_Q is the total number of spectra in that class, $\bar{\gamma}(t_r)$ and $\bar{\sigma}_T(t_r)$ are the relative mean square error on the emissivity estimates (defined in eq. (27)) and the root mean square error on the temperature estimates averaged over the elements of that class, respectively.

Several experiments are carried out by using D-SBTES, Pol-SBTES-1 and by varying the SNR and the rank K of the adopted basis matrix.

Before starting with the analysis of the simulation results, we give insights about the relationship between the specific value of SNR and the impact of noise on the at-sensor radiance. To this purpose, let us consider the spectrum ε_r adopted in the examples presented in Section IV.B. In Figure 10 we compare the spectral at-sensor radiance (\mathbf{L}) for a hypothetical surface with emissivity ε_r and temperature $t_r = 30^\circ\text{C}$, with its noisy version ($\mathbf{L}^{(N)}$) when the noise realizations are generated by setting SNR to 25 dB, 30 dB and 40 dB, respectively.

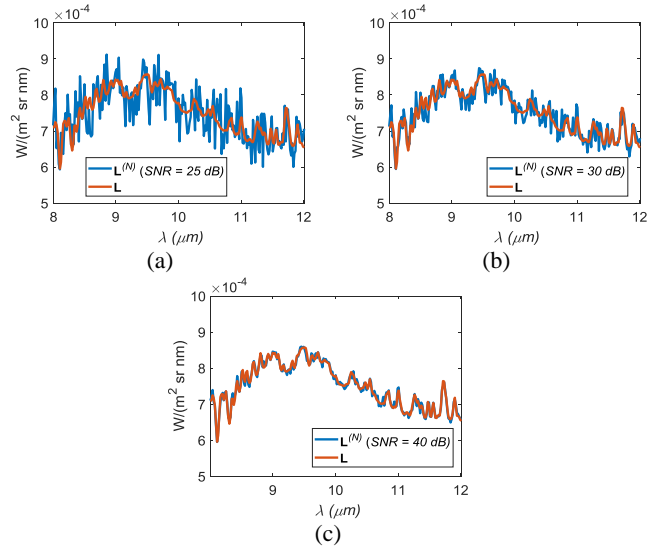


Figure 10. Comparison between the at-sensor radiance spectrum \mathbf{L} and its noisy version $\mathbf{L}^{(N)}$ obtained by adding Gaussian noise with $SNR = 25$ dB (a), $SNR = 30$ dB (b) and $SNR = 40$ dB (c).

Figure 10 a)-c) show that noise significantly impacts on the at-sensor radiance when $SNR = 25$ dB, noise impact is moderate with $SNR = 30$ dB and becomes negligible for $SNR = 40$ dB.

As to the discussion on the performance of the algorithms, let us start by showing the results obtained by the D-SBTES algorithm on the high emissivity spectra. In Figure 11 we plot the values of $\bar{\sigma}_T(t_r)$ (Figure 11 a)) and $\bar{\gamma}(t_r)$ (in percent, Figure 11 b)) obtained by setting $t_r = 30^\circ\text{C}$ and by considering four different values of SNR and three different values of K . As expected, the performance improves when SNR increases, i.e. both $\bar{\sigma}_T(t_r)$ and $\bar{\gamma}(t_r)$ decrease by augmenting SNR . Focusing on the results obtained with $K = 8$, $\bar{\sigma}_T(t_r)$ ranges from a value of about 2 °C for $SNR = 30$ dB to 0.5 °C for $SNR = 45$ dB. Similarly, $\bar{\gamma}(t_r)$ is about 4.3% at $SNR = 30$ dB and reduces to 1% when $SNR = 45$ dB.

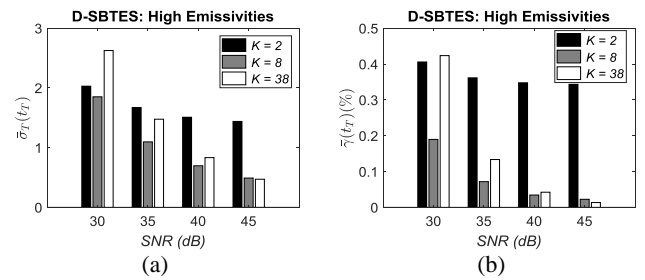


Figure 11. D-SBTES performance on high emissivity spectra for SNR ranging from 30 to 45 dB and $K = 2, 8, 38$. a) Values of $\bar{\sigma}_T(t_r)$; b) values of $\bar{\gamma}(t_r)$ in percent.

For $K = 2$ (low rank basis matrix) both the performance indexes do not change significantly by reducing the noise. This is due to the fact that the basis matrix with rank $K = 2$ does not properly address the emissivity subspace, and the performance is almost entirely influenced by the high approximation error $\Delta\varepsilon$.

Results in Figure 11 also show that, when noise is not negligible (low SNR), by augmenting the rank of the basis matrix and

reducing the approximation error ($\Delta \epsilon$), the performance of the algorithm does not necessarily improve. Focusing on $SNR=30\text{ dB}$, we note that both $\bar{\sigma}_\tau(t_\tau)$ and $\bar{\gamma}(t_\tau)$ reduce their values when we switch from $K=2$ to $K=8$. However, by further increasing K to 38, the values of both the indexes grow. This confirms the conclusion drawn in Section IV.B that in presence of noise the performance of a generic SBTES algorithm degrades by increasing the rank of the basis matrix. Of course, such an effect vanishes when the noise is negligible. In fact, with $SNR=45\text{ dB}$, both $\bar{\sigma}_\tau(t_\tau)$ and $\bar{\gamma}(t_\tau)$ have decreasing trend for increasing values of K . Similar conclusions can be drawn for the Pol-SBTES-1 algorithm whose performance indexes, evaluated on high emissivity spectra and with $t_\tau=30\text{ }^\circ\text{C}$, are shown for completeness in Figure 12.

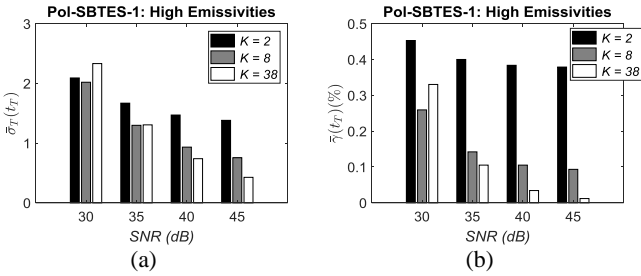


Figure 12. Pol-SBTES-1 performance on high emissivity spectra for SNR ranging from 30 to 45 dB and $K=2,8,38$. a) Values of $\bar{\sigma}_\tau(t_\tau)$; b) values of $\bar{\gamma}(t_\tau)$ in percent.

According to the general discussion in Section IV.B, it is expected that both the algorithms attain the worst performance on the lowest emissivity spectra. Experimental evidence is given in Figure 13 where we plot the values of $\bar{\sigma}_\tau(t_\tau)$ (Figure 13 a) and $\bar{\gamma}(t_\tau)$ (Figure 13 b)) obtained by D-SBTES and Pol-SBTES-1 on simulated data regarding the three classes of spectra “High”, “Low” and “Very Low”. Specifically, the presented results were obtained by setting $SNR=35\text{ dB}$ and $K=8$.

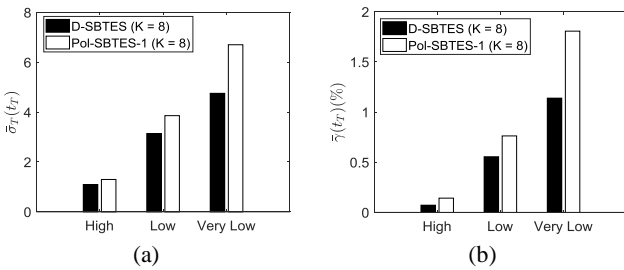


Figure 13. Estimation performance of D-SBTES and Pol-SBTES-1 obtained on the High, Low and Very Low classes of emissivity spectra when SNR is set to 35 dB and $K=8$. (a) Values of $\bar{\sigma}_\tau(t_\tau)$; (b) values of $\bar{\gamma}(t_\tau)$ in percent.

Figure 13 a) and b) clearly show that, regardless of the adopted algorithm, the lowest values of both the performance indexes are obtained for the class of high emissivity spectra, and the highest values pertain the class of very low emissivity spectra. Furthermore, results in Figure 13 show that the performance of a given SBTES algorithm strictly depends on the specific

structure of the adopted basis matrix. In fact, with the same K we obtain different values of the performance indexes by applying D-SBTES and Pol-SBTES-1. Specifically, with $K=8$ and $SNR=35\text{ dB}$, D-SBTES outperforms Pol-SBTES-1 for all the classes of emissivity spectra.

To deepen the comparison between the two algorithms a further experiment is discussed. In such an experiment, we compare the two algorithms when the corresponding basis matrix allows low approximation errors to be obtained on the considered data set. For this purpose, the rank K of the adopted basis matrix is chosen so as to attain a relative approximation error $\frac{\|\Delta \epsilon\|}{\|e\|}$ less than 2% for all the spectra of the \mathbf{V}^{TOT} dictionary. For D-SBTES, the minimum value of K that meets such requirement is 8, whereas for Pol-SBTES-1 is 24. In Figure 14 we plot the values of $\bar{\sigma}_\tau(t_\tau)$ (Figure 14 a) and $\bar{\gamma}(t_\tau)$ (Figure 14 b)) obtained on the class of High emissivity spectra by the two algorithms and for SNR ranging from 20 to 50 dB.

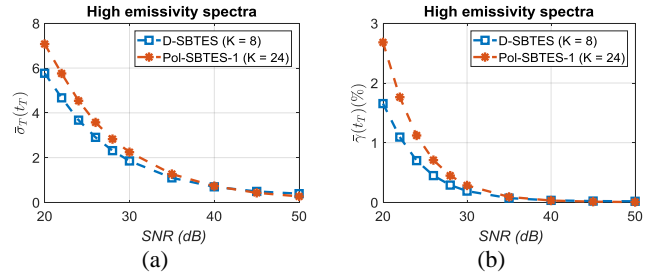


Figure 14. Estimation performance of D-SBTES and Pol-SBTES-1 obtained on the High emissivity spectra with $SNR=20, \dots, 50\text{ dB}$. (a) Values of $\bar{\sigma}_\tau(t_\tau)$; (b) values of $\bar{\gamma}(t_\tau)$ in percent.

Figure 14 shows that D-SBTES outperforms Pol-SBTES-1 for low values of SNR , where noise effects are not negligible with respect to those of the emissivity approximation error $\Delta \epsilon$. This is basically due to the fact that D-SBTES makes use of a basis matrix with rank much lower than that adopted by Pol-SBTES-1. For high SNR , the two algorithms have quite similar performance. This is not surprising because for high SNR , the performance are basically determined by $\Delta \epsilon$, and, in this experiments, the ranks (K) of the basis matrices adopted by the two algorithms are chosen so as to have approximation errors in the same order of magnitude. Similar conclusions can be drawn by considering the class of low emissivity spectra and that of very low emissivity spectra. For completeness, in Figure 15 we show the values of the performance indexes obtained for these two classes of emissivity spectra. As expected, the performance of both the algorithms decreases when the spectral emissivity decreases. Figure 15 confirms that D-SBTES provides the best performance.

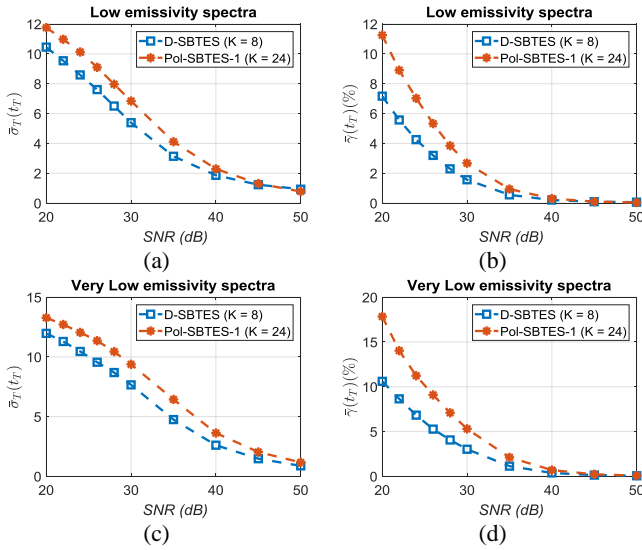


Figure 15. Estimation performance of D-SBTES and Pol-SBTES-1 with $SNR=20, \dots, 50$ dB obtained on: (a)-(b) Low emissivity spectra and (c)-(d) Very Low emissivity spectra.

Further experiments are carried out in order to show the dependence of the algorithms performance on the specific value of the surface temperature. For this purpose, three temperatures are considered: 20°C , 30°C and 40°C . In Figure 16 we plot the performance indexes for the D-SBTES algorithm for different SNR and for the three considered temperatures. Results refer to the class of high emissivity spectra and were obtained by setting $K=8$. Similarly, in Figure 17 we present the results obtained by Pol-SBTES-1 with $K=24$.

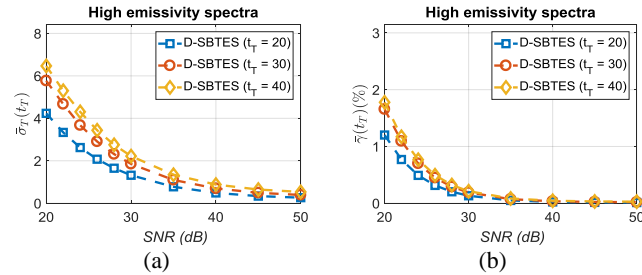


Figure 16. Estimation performance of D-SBTES obtained, by setting $K=8$, on High emissivity spectra with $SNR=20, \dots, 50$ dB and surface temperatures of 20°C , 30°C and 40°C : (a) values of $\bar{\sigma}_T(t_T)$; (b) values of $\bar{\gamma}(t_T)$ in percent.

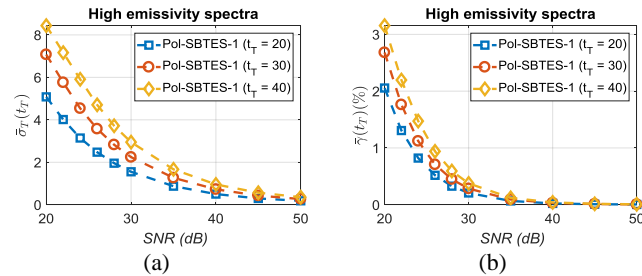


Figure 17. Estimation performance of Pol-SBTES-1 obtained, by setting $K=24$, on High emissivity spectra with $SNR=20, \dots, 50$ dB and surface temperatures of 20°C , 30°C and 40°C : (a) values of $\bar{\sigma}_T(t_T)$; (b) values of $\bar{\gamma}(t_T)$ in percent.

Results in Figure 16 and Figure 17 show that both the algorithms provide better performance with low temperatures. This is particularly true for low SNR . The effects of the surface temperature vanishes when noise is negligible. Very similar results (not reported here) were obtained by applying the algorithms to low and very low emissivity spectra.

To conclude this section we give examples concerning the results provided by the two algorithms in terms of estimated temperature and estimated spectral emissivity. Specifically, we considered three spectra of the dictionary having very low emissivity ($\epsilon_1, \rho_{\epsilon_1} = 0.08$), low emissivity ($\epsilon_2, \rho_{\epsilon_2} = 0.41$) and high emissivity ($\epsilon_3, \rho_{\epsilon_3} = 0.9$). For each of them we simulated the noisy at-sensor radiance assuming $t_T = 30^\circ\text{C}$ and we applied both D-SBTES with $K=8$ and Pol-SBTES-1 with $K=24$. In Figure 18 we show the results obtained when SNR was set to 30 dB. Specifically, in Figure 18 a) we plot the temperature estimates provided by the two algorithms whereas in Figure 18 b)-d) we plot the estimates of the spectral emissivity for the three considered spectra. As expected the worst performance is obtained by applying both the algorithms to ϵ_1 . Conversely, the best performance in terms of estimated temperature and retrieved spectrum is obtained for the high emissivity spectrum ϵ_3 .

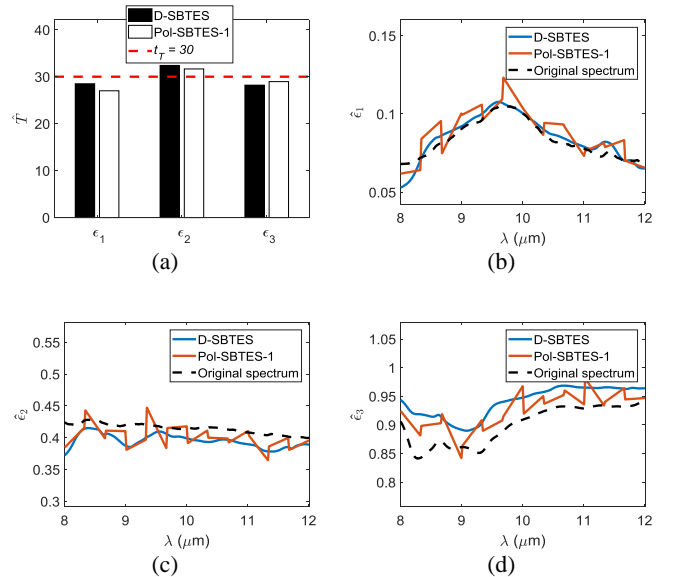


Figure 18. Results obtained by applying the D-SBTES and the Pol-SBTES-1 algorithms to three simulated noisy radiances with $SNR=30$ dB. (a) Temperatures estimates; (b) emissivity estimates on the very low emissivity spectrum; (c) emissivity estimates on the low emissivity spectrum; (d) emissivity estimates on the high emissivity spectrum. D-SBTES was applied by setting $K=8$ and Pol-SBTES-1 was applied by setting $K=24$.

The performance of both the algorithms considerably improves when high SNR is considered. It is exemplified in Figure 19 a)-d) that show the results obtained by applying the two algorithms to data simulated with $SNR=50$ dB.

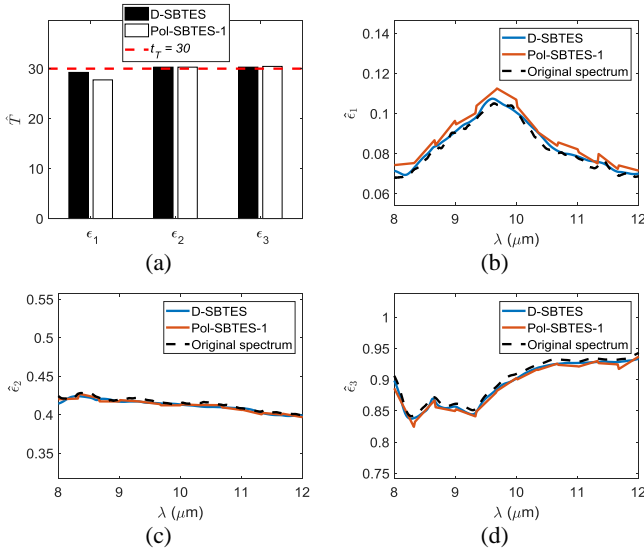


Figure 19. Results obtained by applying the D-SBTES and the Pol-SBTES-1 algorithms to three simulated noisy radiances with $SNR = 50$ dB . (a) Temperatures estimates; (b) emissivity estimates on the very low emissivity spectrum; (c) emissivity estimates on the low emissivity spectrum; (d) emissivity estimates on the high emissivity spectrum. D-SBTES was applied by setting $K = 8$ and Pol-SBTES-1 was applied by setting $K = 24$.

V. CONCLUSIONS

In this paper a class of TES algorithms, named SBTES, was derived by exploiting the assumption that the emissivity spectra of natural and man-made materials can be well represented in a given subspace of the original data space. Specifically, by combining the subspace model with the Gaussian model for the noise affecting LWIR hyperspectral data, the ML estimators of both the surface temperature and the spectral emissivity were derived. The SBTES approach originates several TES algorithms whose specific form depends on the particular basis matrix adopted to address the emissivity subspace. In this paper, two algorithms were presented. The first, named Pol-SBTES- p , was obtained by assuming that the emissivity spectrum can be approximated as a piecewise polynomial function (of degree p) of the sensor central wavelengths. Such a method does not require *a-priori* information about the materials in the considered data set. The second proposed algorithm, named D-SBTES, exploits the availability of a given dictionary of emissivity spectra to estimate the subspace basis matrix. Pol-SBTES- p and D-SBTES are not the only algorithms that can be derived within from the SBTES general framework. Many other algorithms can be obtained by using different basis vectors such as, for instance, those defining the Discrete Cosine Transform or those derived from a given discrete wavelet mother function.

Issues related to the noise and the subspace approximation errors were analyzed in relation to the performance of SBTES algorithms. Such an analysis was carried out by using theoretical bounds on the accuracy of the temperature and emissivity estimators and experimental results obtained on simulated data. The analysis led to several general conclusions.

Firstly, it revealed that the basis matrix adopted by the

SBTES algorithm must be designed so as to meet two conflicting requirements. On one hand, it should have high rank in order to reduce the subspace approximation error as much as possible. On the other hand, it should have low rank in order to reduce the detrimental impact of the noise on the estimation performance.

Secondly, the analysis showed that, regardless of the specific adopted algorithm, the worst performance is obtained on the materials having low spectral emissivity. Furthermore, in presence of noise, we obtain the best performance for the materials with the lowest surface temperature.

Those general conclusions are in compliance with the results obtained by applying the Pol-SBTES-1 and the D-SBTES algorithms to simulated data. The comparison between the two algorithms showed that D-SBTES gives the best overall performance. This is basically due to the fact that dictionary based basis estimation yields compact and accurate representation basis. Of course, it is expected to work well on those materials whose spectral emissivity features are sufficiently represented in the dictionary that originates the basis matrix. Anomalous (with respect to the dictionary) spectral emissivities may result in very poor performance. For example, think of an object (e.g. a specific gas) having a narrow band spectral feature not present in the dictionary atoms.

It is worth noting that, the proposed analysis does not claim to be exhaustive. Several open issues still remain. For instance, it could be interesting to study how performance of SBTES algorithms depends on the specific values of the atmospheric model parameters such as the water vapor concentration or the atmospheric temperature profile. It would also be interesting to analyze the impact of the errors of the atmospheric compensation on the estimation performance of the SBTES algorithms. The issues to analyze are too many to enclose them in a single paper. However, the methodological approach presented in this work is surely suitable to extend the analysis of the SBTES algorithms to other issues and is our intention to face with some of them in the near future.

We conclude by remarking that, although it would be interesting to extend our analysis to real LWIR hyperspectral data, we are unable to do so because of the lack of appropriate experimental data sets. In fact, such an analysis would require a great amount of atmospheric compensated hyperspectral images equipped with precise measurements of the surface temperature and emissivity spectra performed on a per-pixel basis. Such extensive data set is not available and simulation is the only tool to perform accurate sensitivity analysis and get reliable conclusions in realistic scenarios while having the control and full knowledge of each important parameter.

APPENDIX I

In this appendix we derive the expression of the ML estimators of t (\hat{t}) and \mathbf{a} ($\hat{\mathbf{A}}$) reported in eq. (5) and eq. (8). According to the multivariate Gaussian assumption and the subspace model in eq. (4), the log-likelihood function for $\mathbf{Y}(t, \mathbf{e})$ can be written as:

$$\log f(\mathbf{Y}; t, \mathbf{a}) = u - \frac{1}{2} \cdot \left\| \Gamma_Y^{-1/2} \cdot \{\mathbf{Y} - \bar{\mathbf{D}}(t) \cdot \mathbf{U} \cdot \mathbf{a}\} \right\|^2 \quad (\text{I-1})$$

where, $\Gamma_Y = \mathbf{D}_r^{-1} \cdot \Gamma_N \cdot \mathbf{D}_r^{-1} = \mathbf{D}_r^{-2} \cdot \Gamma_N$ and u is a constant depending on the determinant of Γ_Y . The maximum likelihood estimates of t and \mathbf{a} are obtained by solving the following optimization problem ([10]):

$$\begin{aligned} \{\hat{T}, \hat{\mathbf{A}}\} &= \arg \max_{t, \mathbf{a}} \{\log f(\mathbf{Y}; t, \mathbf{a})\} \\ &= \arg \min_{t, \mathbf{a}} \{\Theta(t, \mathbf{a})\} \\ \Theta(t, \mathbf{a}) &= \left\| \Gamma_Y^{-1/2} \cdot \{\mathbf{Y} - \bar{\mathbf{D}}(t) \cdot \mathbf{U} \cdot \mathbf{a}\} \right\|^2 \end{aligned} \quad (\text{I-2})$$

Taking the derivatives of $\Theta(t, \mathbf{a})$ with respect to \mathbf{a} and equating it to zero we derive ([10]):

$$\mathbf{A}(t) = \tilde{\mathbf{U}}(t)^+ \cdot \tilde{\mathbf{Y}} \quad (\text{I-3})$$

with

$$\tilde{\mathbf{U}}(t) = \Gamma_Y^{-1/2} \cdot \bar{\mathbf{D}}(t) \cdot \mathbf{U} = \Gamma_N^{-1/2} \cdot \mathbf{D}_r \cdot \bar{\mathbf{D}}(t) \cdot \mathbf{U} \quad (\text{I-4})$$

and $\tilde{\mathbf{Y}} = \Gamma_N^{-1/2} \cdot \mathbf{D}_r \cdot \mathbf{Y}$. In eq. (I-3) the super-script “+” denotes the Moore–Penrose pseudo inverse. Replacing \mathbf{a} with $\mathbf{A}(t)$ in the expression of $\Theta(t, \mathbf{a})$ we can write:

$$\begin{aligned} \Psi(t) &= \min_{\mathbf{a}} \{\Theta(t, \mathbf{a})\} = \Theta(t, \mathbf{A}(t)) = \\ &= \left\| \tilde{\mathbf{Y}} - \tilde{\mathbf{U}}(t) \cdot \tilde{\mathbf{U}}(t)^+ \cdot \tilde{\mathbf{Y}} \right\|^2 \end{aligned} \quad (\text{I-5})$$

Denoting as $\mathbf{P}(t) = \tilde{\mathbf{U}}(t) \cdot \tilde{\mathbf{U}}(t)^+$ the projection matrix on the subspace spanned by $\tilde{\mathbf{U}}(t)$ and as $\mathbf{P}^\perp(t) = \mathbf{I} - \tilde{\mathbf{U}}(t) \cdot \tilde{\mathbf{U}}(t)^+$ (\mathbf{I} is $N_b \times N_b$ identity matrix) the projection matrix on its orthogonal complement, eq. (I-4) can be rewritten as:

$$\Psi(t) = \left\| \mathbf{P}^\perp(t) \cdot \tilde{\mathbf{Y}} \right\|^2 \quad (\text{I-5})$$

Now, we can obtain \hat{T} as:

$$\hat{T} = \arg \min_t \{\Theta(t, \mathbf{A}(t))\} = \arg \min_t \Psi(t) \quad (\text{I-6})$$

and:

$$\hat{\mathbf{A}} = \mathbf{A}(\hat{T}) = \tilde{\mathbf{U}}(\hat{T})^+ \cdot \tilde{\mathbf{Y}} \quad (\text{I-3})$$

APPENDIX II

In this appendix, we first derive the Fisher information matrix and its inverse for the unbiased estimators of the deterministic parameters t and \mathbf{a} in the subspace based estimation problem defined by eqs. (3) and (4). Then, exploiting these results, we obtain the expression of the Cramer-Rao Lower Bound (CRLB) for the unbiased estimators of the surface temperature t (in absence of emissivity approximation error $\Delta\varepsilon = 0$), and the lower bound for the emissivity estimation error.

According to the model introduced in Section II, the observation vector \mathbf{Y} is Gaussian distributed with mean vector $\boldsymbol{\mu}_Y(t, \mathbf{a})$ and covariance matrix Γ_Y :

$$\begin{aligned} \boldsymbol{\mu}_Y(t, \mathbf{a}) &= \bar{\mathbf{D}}(t) \cdot \mathbf{U} \cdot \mathbf{a} \\ \Gamma_Y &= \mathbf{D}_r^{-1} \cdot \Gamma_N \cdot \mathbf{D}_r^{-1} \\ \bar{\mathbf{D}}(t) &= \mathbf{D}_{BB}(t) - \mathbf{D}_L \end{aligned} \quad (\text{II-1})$$

In the case of Gaussian observations the Fisher information matrix $\mathbf{F}_Y(t, \mathbf{a})$ is given by ([10], pag. 47):

$$\mathbf{F}_Y(t, \mathbf{a}) = \mathbf{A}(t, \mathbf{a})^T \cdot \Gamma_Y^{-1} \cdot \mathbf{A}(t, \mathbf{a}) \quad (\text{II-2})$$

where $\mathbf{A}(t, \mathbf{a})$ is the $N_b \times K + 1$ matrix containing the partial derivatives of the mean vector with respect to the parameters t and \mathbf{a} , i.e.:

$$\mathbf{A}(t, \mathbf{a}) = \begin{bmatrix} \frac{\partial}{\partial t} \boldsymbol{\mu}_Y(t, \mathbf{a}), \frac{\partial}{\partial \mathbf{a}} \boldsymbol{\mu}_Y(t, \mathbf{a}) \end{bmatrix} \quad (\text{II-3})$$

In our case, it can be easily proved that:

$$\begin{aligned} \frac{\partial}{\partial t} \boldsymbol{\mu}_Y(t, \mathbf{a}) &= \mathbf{D}_{BB}^{(l)}(t) \cdot \mathbf{U} \cdot \mathbf{a} \\ \frac{\partial}{\partial \mathbf{a}} \boldsymbol{\mu}_Y(t, \mathbf{a}) &= \bar{\mathbf{D}}(t) \cdot \mathbf{U} \end{aligned} \quad (\text{II-4})$$

where $\mathbf{D}_{BB}^{(l)}(t)$ is the diagonal matrix whose l -th entry $d_{BB}^{(l)}(\lambda_l, t)$ is the derivative of the Plank’s law with respect to the temperature at the central wavelength λ_l .

Combining eqs. (II-3) and (II-4) and recalling that the covariance matrix Γ_Y is symmetric, $\mathbf{F}_Y(t, \mathbf{a})$ can be rewritten as:

$$\begin{aligned} \mathbf{F}_Y(t, \mathbf{a}) &= \left[\bar{\mathbf{D}}^{(l)}(t) \cdot \mathbf{U} \cdot \mathbf{a}, \tilde{\mathbf{U}}(t) \right]^T \cdot \left[\bar{\mathbf{D}}^{(l)}(t) \cdot \mathbf{U} \cdot \mathbf{a}, \tilde{\mathbf{U}}(t) \right] \\ \bar{\mathbf{D}}^{(l)}(t) &= \Gamma_Y^{-1/2} \cdot \mathbf{D}_{BB}^{(l)}(t) = \Gamma_N^{-1/2} \cdot \mathbf{D}_r \cdot \mathbf{D}_{BB}^{(l)}(t) \\ \tilde{\mathbf{U}}(t) &= \Gamma_Y^{-1/2} \cdot \bar{\mathbf{D}}(t) \cdot \mathbf{U} = \Gamma_N^{-1/2} \cdot \mathbf{D}_r \cdot \bar{\mathbf{D}}(t) \cdot \mathbf{U} \end{aligned} \quad (\text{II-5})$$

After some algebraic calculation we obtain $\mathbf{F}_Y(t, \mathbf{a})$ in the following block form:

$$\mathbf{F}_Y(t, \mathbf{a}) = \begin{bmatrix} c & \mathbf{b}^T \\ \mathbf{b} & \mathbf{R} \end{bmatrix} \quad (\text{II-6})$$

where $c = \mathbf{a}^T \cdot \mathbf{U}^T \cdot \bar{\mathbf{D}}^{(l)}(t) \cdot \bar{\mathbf{D}}^{(l)}(t) \cdot \mathbf{U} \cdot \mathbf{a}$ is a scalar quantity, $\mathbf{b} = \tilde{\mathbf{U}}(t)^T \cdot \bar{\mathbf{D}}^{(l)}(t) \cdot \mathbf{U} \cdot \mathbf{a}$ is a $K \times 1$ vector and $\mathbf{R} = \tilde{\mathbf{U}}(t)^T \cdot \tilde{\mathbf{U}}(t)$ is a $K \times K$ invertible matrix. According to the block matrix inversion formula:

$$\mathbf{F}_Y^{-1}(t, \mathbf{a}) = \begin{bmatrix} \frac{1}{\beta} & -\frac{1}{\beta} \mathbf{b}^T \cdot \mathbf{R}^{-1} \\ -\frac{1}{\beta} \mathbf{R}^{-1} \cdot \mathbf{b} & \mathbf{F}_a^{-1}(t, \mathbf{a}) \end{bmatrix} \quad (\text{II-7})$$

with:

$$\beta = c - \mathbf{b}^T \cdot \mathbf{R}^{-1} \cdot \mathbf{b} \quad (\text{II-8})$$

and

$$\mathbf{F}_a^{-1}(t, \mathbf{a}) = \mathbf{R}^{-1} + \frac{1}{\beta} \mathbf{R}^{-1} \cdot \mathbf{b} \cdot \mathbf{b}^T \cdot \mathbf{R}^{-1} \quad (\text{II-9})$$

The first element on the diagonal of $\mathbf{F}_Y^{-1}(t, \mathbf{a})$ is the CRLB for the unbiased estimators of the parameter t ($CRLB_t(t, \mathbf{a})$) ([10]). Accounting for the expressions of c , \mathbf{b} and \mathbf{R} , it can be written as:

$$CRLB_i(t, \mathbf{a}) = \frac{1}{\|\mathbf{P}^\perp(t) \cdot \bar{\mathbf{D}}^{(i)}(t) \cdot \mathbf{U} \cdot \mathbf{a}\|^2} \quad (\text{II-10})$$

$$\mathbf{P}^\perp(t) = \mathbf{I} - \tilde{\mathbf{U}}(t) \cdot [\tilde{\mathbf{U}}(t)^T \cdot \tilde{\mathbf{U}}(t)]^{-1} \cdot \tilde{\mathbf{U}}(t)^T$$

The remaining diagonal entries of $\mathbf{F}_v^{-1}(t, \mathbf{a})$ (i.e. those of $\mathbf{F}_a^{-1}(t, \mathbf{a})$) are the CRLBs for the unbiased estimators of the elements of \mathbf{a} .

In order to derive a lower bound for the estimation error on the emissivity vector, let us start by defining the mean square estimation error as:

$$\mathcal{G}(t, \mathbf{a}) = E\left[\|\hat{\boldsymbol{\varepsilon}} - \boldsymbol{\varepsilon}\|^2\right] \quad (\text{II-11})$$

with E denoting the statistical expectation operator.

According to the adopted subspace model:

$$\begin{aligned} \mathcal{G}(t, \mathbf{a}) &= \text{tr}\left\{E\left[\mathbf{U} \cdot (\hat{\mathbf{A}} - \mathbf{a}) \cdot (\hat{\mathbf{A}} - \mathbf{a})^T \cdot \mathbf{U}^T\right]\right\} = \\ &= \text{tr}\left\{\mathbf{U} \cdot E\left[(\hat{\mathbf{A}} - \mathbf{a}) \cdot (\hat{\mathbf{A}} - \mathbf{a})^T\right] \cdot \mathbf{U}^T\right\} \end{aligned} \quad (\text{II-12})$$

where tr denotes the trace of a matrix. For the generic unbiased estimator of \mathbf{a} , the statistical expectation in eq. (II-12) coincides with its covariance matrix $\mathbf{C}_{\hat{\mathbf{A}}}$, for which the following inequality holds ([10]):

$$\mathbf{C}_{\hat{\mathbf{A}}} - \mathbf{F}_a^{-1}(t, \mathbf{a}) \geq \mathbf{0} \quad (\text{II-13})$$

where $\geq \mathbf{0}$ is to be interpreted as positive semidefinite. As a consequence we can write:

$$\mathbf{U} \cdot \mathbf{C}_{\hat{\mathbf{A}}} \cdot \mathbf{U}^T - \mathbf{U} \cdot \mathbf{F}_a^{-1}(t, \mathbf{a}) \cdot \mathbf{U}^T \geq \mathbf{0} \quad (\text{II-14})$$

Now combining eqs. (II-12) and (II-14), we obtain:

$$\begin{aligned} \mathcal{G}(t, \mathbf{a}) &\geq \mathcal{G}_{LB}(t, \mathbf{a}) \\ \mathcal{G}_{LB}(t, \mathbf{a}) &= \text{tr}\left\{\mathbf{U} \cdot \mathbf{F}_a^{-1}(t, \mathbf{a}) \cdot \mathbf{U}^T\right\} \end{aligned} \quad (\text{II-15})$$

that applies to any unbiased estimator of $\boldsymbol{\varepsilon} = \mathbf{U} \cdot \mathbf{a}$.

Recalling eq. (II-9) and the specific expressions of \mathbf{R} , \mathbf{b} and β , the lower bound in eq. (II-15), is rewritten as:

$$\begin{aligned} \mathcal{G}_{LB}(t, \mathbf{a}) &= \text{tr}\left\{\mathbf{U} \cdot [\tilde{\mathbf{U}}(t)^T \cdot \tilde{\mathbf{U}}(t)]^{-1} \cdot \mathbf{U}^T\right\} + \\ &+ CRLB_i(t, \mathbf{a}) \cdot \|\mathbf{U} \cdot \tilde{\mathbf{U}}(t) \cdot \bar{\mathbf{D}}^{(i)}(t) \cdot \mathbf{U} \cdot \mathbf{a}\|^2 \end{aligned} \quad (\text{II-16})$$

REFERENCES

- [1]. M. T. Eismann, *Hyperspectral Remote Sensing*. SPIE Press, Bellingman, Washington, 2012.
- [2]. D. G. Manolakis, S. E. Golowich, R. S. Di Pietro, "Long-Wave Infrared Hyperspectral Remote Sensing of Chemical Clouds," *IEEE Signal Processing Magazine*, Vol. 31, No. 4, pp. 120-141, July 2014.
- [3]. D. Gu, A. R. Gillespie, A.B. Kahle, and F. D. Palluconi, "Autonomous Atmospheric Compensation (AAC) of High Resolution Hyperspectral Thermal Infrared Remote-Sensing Imagery," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 38, No. 6, pp. 2557-2570, Nov. 2000.
- [4]. S. J. Young, B. R. Johnson, and J. A. Hackwell, "An in-scene method for atmospheric compensation of thermal hyperspectral data," *Journal of Geophysical Research: Atmospheres*, Vol. 107, No. D24, pp. 14-20, 2002.

- [5]. A. Gillespie, S. Rokugawa., T. Matsunaga, S. Cothorn, S. Hook, A. B. Kahle, "A Temperature and Emissivity Separation Algorithm for Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) Images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 36, no. 4, pp. 1113-1126, 1998.
- [6]. M. Pivovarník, S. J. S. Khalsa., J. C. Jiménez-Muñoz, F. Zemek, , "Improved Temperature and Emissivity Separation Algorithm for Multispectral and Hyperspectral Sensors," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 4, pp. 1944 - 1953, 2017
- [7]. J. Cheng, S. Liang, J. Wang, X. Li, "A Stepwise Refining Algorithm of Temperature and Emissivity Separation for Hyperspectral Thermal Infrared Data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 3, pp. 1588 - 1597, 2010.
- [8]. N. Wang, H. Wu, F. Nerry, C. Li, Z. L. Li, "Temperature and Emissivity Retrievals from Hyperspectral Thermal Infrared Data using Linear Spectral Emissivity Constraint," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 4, pp. 1291 - 1303, 2011.
- [9]. J. R. Schott, *Remote Sensing : The Image Chain Approach: The Image Chain Approach*, 2nd ed. New York, NY: Oxford University Press, 2007.
- [10]. S.M. Kay, *Fundamental of statistical signal processing: Estimation Theory*, Prentice-Hall, 1998.
- [11]. J. Meola, M. T. Eismann, R. L. Moses, and J. N. Ash, "Modeling and estimation of signal-dependent noise in hyperspectral imagery," *Applied Optics*, Vol. 50, No. 21, pp. 3829-3846, 2011.
- [12]. M.L. Uss, B. Vozel, V.V. Lukin, K. Chehdi, "Local Signal-Dependent Noise Variance Estimation From Hyperspectral Textural Images," *IEEE Journal of Selected Topics in Signal Processing*, Vol. 5, No.3, pp. 459-486, June 2011.
- [13]. N. Acito, M. Diani, G. Corsini, "Subspace-Based Striping Noise Reduction in Hyperspectral Images, *IEEE Transactions on Geoscience and Remote Sensing*", Vol.49, No. 4, pp. 1325-1342, February 2011.
- [14]. W. C. Therrien, *Discrete Random Signals and Statistical Signal Processing*, Prentice-Hall Signal Processing Series, April 1992.
- [15]. A.M. Baldridge, S.J. Hook, C.I. Grove, G. Rivera, "The ASTER spectral library version 2.0," *Remote Sens. Environ.*, Vol. 113, No. 5, pp. 711-715, April 2009.
- [16]. Z. Wan, D. Ng, J. Dozier, "Spectral emissivity measurements of land-surface materials and related radiative transfer simulations," *Advances in Space Research*, Vol. 14, No. 3, pp. 91-94, March 1994.
- [17]. S. Kotthaus, T. E.L. Smith, M. J. Wooster, C.S.B. Grimmond, "Derivation of an urban materials spectral library through emittance and reflectance spectroscopy," *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 94, pp. 194-212, August 2014.
- [18]. J. Meola, M. T. Eismann, R. L. Moses, and J. N. Ash, "Modeling and estimation of signal-dependent noise in hyperspectral imagery," *Applied Optics*, Vol. 50, No. 21, pp. 3829-3846, 2011.
- [19]. M.L. Uss, B. Vozel, V.V. Lukin, K. Chehdi, "Local Signal-Dependent Noise Variance Estimation From Hyperspectral Textural Images," *IEEE Journal of Selected Topics in Signal Processing*, Vol. 5, No.3, pp. 459-486, June 2011.
- [20]. N. Acito, M. Diani, G. Corsini, "Signal-Dependent Noise Modeling and Model Parameter Estimation in Hyperspectral Images," *IEEE Trans. on Geos. and Rem. Sensing*, Vol. 49, No. 8, pp. 2957-2971, August 2011.
- [21]. W. R. Johnson, S. J. Hook, P. Mouroulis, D. W. Wilson, S. D. Gunapala, V. Realmuto, A. Lamborn, C. Paine, J. M. Mumolo, B. T. Eng, "HyTES: Thermal Imaging Spectrometer Development," *IEEE 2011 Aerospace Conference*, pp. 1-8, 2011.
- [22]. T. Skauli, "Sensor noise informed representation of hyperspectral data, with benefits for image storage and processing," *Optics Express*, Vol. 19, No. 14, pp. 13031-13046, June 2011.