The role of spatial information in disentangling the irradiance-reflectance-transmittance ambiguity

Sandra Jiménez and Jesús Malo

Abstract—In the satellite hyperspectral measures the contributions of light, surface, and atmosphere are mixed. Applications need separate access to the sources. Conventional inversion techniques usually take a pixel-wise, spectral-only approach. However, recent improvements in retrieving surface and atmosphere characteristics use heuristic spatial smoothness constraints.

In this paper we theoretically justify such heuristics by analyzing the impact of spatial information on the uncertainty of the solution. The proposed analysis allows to assess in advance the uniqueness (or robustness) of the solution depending on the curvature of a likelihood surface. In situations where pixel-based approaches become unreliable it turns out that the consideration of spatial information always makes the problem to be better conditioned. With the proposed analysis this is easily understood since the curvature is consistent with the complexity of the sources measured in terms of the number of significant eigenvalues (or free parameters in the problem). In agreement with recent results in hyperspectral image coding, spatial correlations in the sources imply that the intrinsic complexity of the spatio-spectral representation of the signal is always lower than its spectral-only counterpart. According to this, the number of free parameters in the spatio-spectral inverse problem is smaller so the spatio-spectral approaches are always better than spectral-only approaches.

Experiments using ensembles of actual reflectance values and realistic MODTRAN irradiance and atmosphere radiance and transmittance values show that the proposed analysis successfully predicts the practical difficulty of the problem and the improved quality of spatio-spectral retrieval.

Index Terms—Inverse Problems in Remote Sensing, Curvature Analysis, Spatial Information, Complexity of Irradiance-Reflectance-Transmittance, Spatio-Spectral PCA.

I. INTRODUCTION

The spectral information of (1) the irradiance at the Earth surface, (2) the surface reflectance, and (3) the atmosphere radiance and transmittance, are mixed in the remote sensing acquisition process given the energy integration at each position of the sensor array [1]. Obtaining the spectrum of the quantities of interest from the measurements is the classical inverse problem arising in remote sensing [2]–[4].

In some cases, the simultaneous retrieval of surface and atmospheric characteristics is done by iterative processes based on look-up-tables [3], [6], while in others the question is posed as a Bayesian estimation or error minimization problem to be solved by gradient descent [7], [8]. However, the above conventional approaches to the problem usually operate pixel-wise by taking a spectral-only approach [5]–[9]. In Bayesian and error minimization approaches to the inverse problem it is apparent that these spectral-only techniques exploit the known spectral smoothness of the sources either explicitly by using Tikhonov-like regularization functionals [3] or implicitly by SVD truncation [8], as done in the general context of inverse problems where strong oscillations in the solutions are not desirable [4].

Recent improvements in the simultaneous estimation of surface and atmosphere characteristics use heuristic spatial smoothness constraints [9]–[11]. Intuitively, estimation errors in neighbor pixels will cancel if appropriate spatial smoothing is included in every iteration of the techniques [11], or included as additional regularization terms in the error functions to be minimized [9], [10]. These improvements are leading to a consensus on the qualitative idea that using spatial information is positive in the retrieval application.

In this work we theoretically justify such idea by explicitly analyzing the impact of spatial information on the uncertainty of the solution. Rather than proposing a new retrieval technique that takes advantage of spatial information in some particular way, the aim of this work is presenting an analysis that quantifies the complexity of the retrieval problem. Here instead of starting from the conventional pixel-based approach (and including the spatial smoothness as a convenient constraint afterwards), we formulate the whole problem in a spatio-spectral fashion by considering 3D patches of the hyperspectral cube of arbitrary spatial size. In this way we develop a single formulation in which the amount of considered spatial information depends on the spatial patch size parameter. The (limit) spectral-only case is obtained when the size of the spatial window is reduced to a single pixel.

In this setting, the intrinsic difficulty of the problem, and hence the expected quality of the results can be theoretically assessed in the standard way: by analyzing the curvature of the likelihood function [3]. This standard procedure has been certainly applied (for other purposes) in spectral-only scenarios [7], however, our explicit spatio-spectral formulation is intended to stress the role of the relative amount of spatial versus spectral information in a natural way.

Moreover, in order to understand why including spatial information simplifies the estimation, we formulate the problem in a transformed domain. In particular, we use the spatio-spectral Principal Component Analysis (PCA) transform to represent the signals according to decorrelated components. Here we will see that the specific shape of the PCA eigenfunctions and the characteristic trend of the eigenvalues of the remote sensing sources in the considered situations (pixel-wise versus spatio-spectral) clearly illustrates the different complexity of the problem in both cases. Therefore, the (technically) convenient effect of spatial information in the
curvature of the likelihood can be interpreted in clear statistical terms. Additionally, the low-frequency content of the most relevant PCA eigenfunctions naturally leads to the preference for smooth solutions with no ad-hoc heuristics. All this is consistent with equivalent findings (the benefits of the joint spatio-spectral approach and the smoothness of significant features) in the literature devoted to efficient encoding of hyperspectral signals [12]–[15].

The analysis proposed here, specifically formulated for the remote sensing scenario, is strongly inspired in the ideas presented in [16] (in the context of the color constancy problem in conventional photography). While the technical extension of the formulation to the remote sensing case is a (convenient but) straightforward exercise, the physical meaning of the results is certainly worth to be explicitly confirmed in the remote sensing scenario. It is not obvious that the statistical behavior of the signals in these different domains have to be necessarily the same given the extremely different spatial resolution. Moreover, while the aim in [16] was the statistical explanation of the organization of sensors observed in biological vision; the interest of a similar analysis is certainly worth to be explicitly confirmed in the remote sensing scenario. It is not obvious that the retrieval problem refers to obtaining the unknowns (light, surface and atmosphere spectral characteristics) from the measurements. This problem is not trivial since the measurements are a sum of terms that depend on the unknowns in such a way that (i) different combinations of the unknowns may give rise to the same measurement, and (ii) the spectral variation of the unknowns is lost in the integration over \( \lambda \). Even if, as usually done [17], [18], one neglects terms with products of reflectance (since \( \rho(\lambda) S(\lambda) \ll 1 \)), which is a reasonable approximation in situations in which you don’t have extremely bright surfaces such as snow or metallic roofs), the above issues (i and ii) are still present in the simplified imaging equation:

\[
y_b^\ast = \sum_\lambda R_\lambda (A^\lambda (\lambda) + \sum_\lambda R_\lambda (E^\lambda (\lambda) S^\lambda (\lambda) T^\lambda (\lambda)) (1 + \rho^\lambda (\lambda) S^\lambda (\lambda) + (\rho^\lambda (\lambda) S^\lambda (\lambda))^2 + \cdots))
\]

The retrieval problem refers to obtaining the unknowns (light, surface and atmosphere spectral characteristics) from the measurements. This problem is not trivial since the measurements are a sum of terms that depend on the unknowns in such a way that (i) different combinations of the unknowns may give rise to the same measurement, and (ii) the spectral variation of the unknowns is lost in the integration over \( \lambda \). Even if, as usually done [17], [18], one neglects terms with products of reflectance (since \( \rho(\lambda) S(\lambda) \ll 1 \)), which is a reasonable approximation in situations in which you don’t have extremely bright surfaces such as snow or metallic roofs), the above issues (i and ii) are still present in the simplified imaging equation:

\[
y_b^\ast \approx \sum_\lambda R_\lambda (A^\lambda (\lambda) + E^\lambda (\lambda) S^\lambda (\lambda) T^\lambda (\lambda) ) \]

Note that Eq. [3] has the classical inverse-problem form [4], [19], which commonly appears in many ill-posed vision problems [20]. Actually, the popular color constancy problem of the human and computer vision communities refers to the second term in the sum of Eq. [3] in which irradiance, reflectance and transmittance are mixed up [20]. In the remote sensing scenario the problem seems worse since, at least, there is one additional term (the one depending on \( A \)) that may introduce additional ambiguities in the solutions. The goal of the work is presenting a theoretical analysis that demonstrates the benefits of spatial information in this inverse problem. Here we focus on the simplified Eq. [3] (the more physically relevant terms in Eq. [2]). However, note that more general situations just imply adding extra terms in the sum which do not change the mathematical nature of the problem. For example, including the linear term on \( \rho(\lambda) \) only introduces an additional loop in the estimation described in Section IV.B.

III. BAYESIAN ESTIMATION IN THE PCA DOMAIN

The key of the proposed analysis is using an efficient representation for each of the unknowns. Efficient in the sense of using small number of parameters to describe each unknown. Then, the idea is analyzing how hard is the Bayesian estimation of this small set of parameters (or degrees of freedom) describing the unknowns.

Efficiency in terms of minimum amount of data to represent a signal is related to transform coding [12]–[15], and naturally...
leads to Principal Components Analysis (PCA) \([21]\). In PCA, sample signals from a source (vectors from a statistical ensemble) are used to estimate the covariance matrix of the process. Then, the eigenfunctions of the covariance define a transform to a domain with optimal energy compaction in a minimum number of decorrelated coefficients.

According to the above, after an off-line training stage (e.g. field measurements prior to the actual satellite mission in which samples of the unknowns are collected to estimate the covariance matrices of \(E^{x}(\lambda), T^{x}(\lambda), A^{x}(\lambda)\) and \(S^{x}(\lambda)\)), the unknowns can be written as a linear combination of spatio-spectral basis functions computed through PCA,

\[
E^{x}(\lambda) = \sum_{i=1}^{m^2} E_{i}^{x}(\lambda) e_{i}^{x} \\
T^{x}(\lambda) = \sum_{j=1}^{m^2} T_{j}^{x}(\lambda) t_{j}^{x} \\
A^{x}(\lambda) = \sum_{z=1}^{m^2} A_{x}^{z}(\lambda) a_{z}^{x} \\
S^{x}(\lambda) = \sum_{k=1}^{m^2} S_{k}^{x}(\lambda) s_{k}^{x} \tag{4}
\]

where the range of the coefficient indices, \((i, j, z, k)\), is determined by the number of considered pixels, \(m^2\) (assuming square patches of size \(m \times m\)), and the total number of sensors (or bands) per pixel, \(B\). Note that the superscript, \(x\), in the coefficients is not strictly required since the actual spatio-spectral dependence is in the basis functions (in capital letters). However, it is convenient when we compare the natural spatio-spectral approach (in cubes with \(m > 1\)), with the spectral only approach (in cubes with \(m = 1\)). In this latter case we will consider as many pixels as in the corresponding spatio-spectral case. In that situation (many pixels in a pixel-wise approach) the location superscript is convenient to point out that relations between different locations will not be considered.

Note also that given the intrinsic low dimensionality of the considered sources, the above expansions can be truncated introducing a limited amount of representation error in the measurements. Therefore, accepting some (arbitrarily small) error, the actual number of unknowns reduces to the selected number of terms \((d = m^2 B)\) in the expansions:

\[
\tilde{E}^{x}(\lambda) = \sum_{i=1}^{d_{c}} E_{i}^{x}(\lambda) e_{i}^{x} \\
\tilde{T}^{x}(\lambda) = \sum_{j=1}^{d_{c}} T_{j}^{x}(\lambda) t_{j}^{x} \\
\tilde{A}^{x}(\lambda) = \sum_{z=1}^{d_{c}} A_{x}^{z}(\lambda) a_{z}^{x} \\
\tilde{S}^{x}(\lambda) = \sum_{k=1}^{d_{c}} S_{k}^{x}(\lambda) s_{k}^{x} \tag{5}
\]

Therefore, the measurements (up to a certain representation error, \(e\)) can be expressed as a (truncated) linear combination of the basis functions:

\[
y_{e}^{b} = \sum_{i=1}^{d_{c}} \sum_{\lambda} R_{b}(\lambda) \cdot E_{i}^{x}(\lambda) a_{i}^{x} + \sum_{i=1}^{d_{c}} \sum_{j=1}^{d_{c}} \sum_{k=1}^{d_{c}} G_{ijk,b}^{e} e_{i}^{x} t_{j}^{x} s_{k}^{x} + e_{b}^{e} \tag{6}
\]

where, the trilinear forms, \(G_{ijk,b}^{e}\), are:

\[
G_{ijk,b}^{e} = \sum_{\lambda} R_{b}(\lambda) E_{i}^{x}(\lambda) T_{j}^{x}(\lambda) S_{k}^{x}(\lambda) \tag{7}
\]

\[A. \text{ Bayesian formulation}\]

Assuming the PCA representation, in Bayesian terms, the estimation of the spectral variables at each location, \((E, T, A, S)\), from a set of measurements, \(y\), reduces to the maximization of the posterior probability of the spatio-spectral coefficients \((e, t, a, s)\) for the scene given the measurements: \(P(e, t, a, s|y)\).

Assuming that the error at sensor pixels, \(e\), is i.i.d zero-mean Gaussian with variance, \(\sigma^2\), the posterior probability can be written as:

\[
P(e, t, a, s|y) = \frac{1}{Z} P(y|e, t, a, s) P(e, t, a, s) \tag{8}
\]

where \(P(e, t, a, s)\) can be expressed as:

\[
f(e, t, a, s) = \sum_{z, \lambda} R_{b}(\lambda) A_{x}^{z}(\lambda) a_{z}^{x} - \sum_{i, j, k} G_{ijk,b}^{e} e_{i}^{x} t_{j}^{x} s_{k}^{x}
\]

where \(Z\) is just a normalization constant. The solution to the estimation problem exists and is unique if the posterior probability has a single maximum in the space of coefficients \((e, t, a, s)\).

The difficulty of the estimation problem (or the robustness of the solution to noise) depends on how well conditioned is the problem, which is related to the curvature of the posterior probability at its maximum. This curvature should depend on (1) the ratio between the number of unknowns and the number of available measurements, and (2) the nature of the problem at hand.

The aim of the paper is illustrating that, at least in optical remote sensing, the above issues favor spatio-spectral approaches in front of the conventional spectral-only approach. Here we will show that the problem is better conditioned (the curvature is increased) when using spatio-spectral representations with regard to the spectral-only (pixel-wise) approaches.

\[B. \text{ Uniqueness of the solution from the curvature of the likelihood}\]

Assuming we have no clue on the possible values of the coefficients, i.e. uniform prior \(P(e, t, a, s)\), the curvature of the posterior is determined by the curvature of the likelihood (or the -more convenient- log-likelihood):

\[
L(y|e, t, a, s) = \log P(y|e, t, a, s) \tag{9}
\]

\[
= \frac{1}{2\sigma^2} \sum_{x=1}^{B} (y_{e}^{b} - f(e, t, a, s))^2
\]

Note that this log-likelihood can be understood as a prediction error between the measured values \(y_{e}^{b}\) and the predicted measurements using a PCA statistical model with parameters \((e, t, a, s)\). Accordingly, this log-likelihood is similar to the error term in the functional to be minimized in other approaches \([9], [10]\). The difference here is that spatial correlation contributes much less in the functional to be minimized in other approaches. Nevertheless, the above results show that spatial correlation is important and that this model leads to a domain with optimal energy compaction in a minimum number of decorrelated coefficients. According to the above, after an off-line training stage (e.g. field measurements prior to the actual satellite mission in which samples of the unknowns are collected to estimate the covariance matrices of \(E^{x}(\lambda), T^{x}(\lambda), A^{x}(\lambda)\) and \(S^{x}(\lambda)\)), the unknowns can be written as a linear combination of spatio-spectral basis functions computed through PCA,
The curvature of the log-likelihood can be analyzed by using the eigen decomposition of its Hessian at the considered point: the eigenvectors determine the principal directions of curvature and the associated eigenvalues are related to the amount of curvature at those directions. At a well defined maximum all the eigenvalues -curvatures- must be non-positive since the log-likelihood has to decrease when departing from the maximum. Large eigenvalues (in absolute value) indicate a sharp peak in the log-likelihood and a better conditioned situation (or more robust solutions). Closer to zero eigenvalues indicate a flat posterior and unreliable solutions. The explicit expression for the Hessian is given in Appendix A.

The likelihood (Eq. 10), the associated Hessian (appendix A), and its eigenvalues represent a convenient way of analyzing the difficulty of the problem since they depend on the critical design parameters in remote sensing: (1) number of sensors per pixel, \( B \), (2) spectral sensitivity of the sensors \( (R_b(\lambda) \) also in the trilinear form), and (3) spatial resolution, or number of pixels in the image plane, \( m^2 \). For certain data variability (captured by the basis functions and by the number of necessary components for a given measurement reconstruction error), the curvatures associated to the likelihood may be computed for different design parameters. The examples in Section V illustrate the possibilities of the analysis of the uniqueness of the solution (or the robustness of the solution) by inspecting the eigenvalues of the Hessian. There, we assume a fixed number of sensors with fixed sensitivity and we just explore the effect of the spatial size of the considered block. Then, the eigenvalues of the Hessian inform us about the benefits of using the spatio-spectral representation versus the spectral-only approach.

IV. DEGREES OF FREEDOM IN THE SPATIO-SPECTRAL AND SPECTRAL-ONLY CASES

In this section we first review the trends of second order statistics of the sources in our problem: the irradiance, the surface reflectance, the atmosphere transmittance and atmosphere radiance in the optical range. Then we consider how many parameters we need to know in order to reproduce the measurements and hence the total number of unknowns in the problem.

Realistic surface reflectance samples for the PCA analysis come from hyperspectral images collected with the airborne AHS whisk-broom line-scanning spectrometer operated by the Inst. Nac. Técnica Aeroesp. [22]. The data were acquired in the framework of the joint European AgriSAR 2006 campaign conducted by the European Space Agency (ESA) [23]. In particular, the analyzed set of images were acquired over the DEMMIN (Durable Environmental Multidisciplinary Monitoring Information Network) test site [24]. The spectrometer, with a FOV of 90°, was operated at a flight altitude of 991 m, resulting in a spatial resolution of 2.5 m at nadir and a swath about 2000 m. AHS covers 20 spectral channels in the VNIR and 43 in the SWIR, however, for the sake of simplicity in the experiments, we limit ourselves to the visible range (450 to 801 nm). In order to obtain ground reflectance data the images were atmospherically corrected using the ATCOR4 model [25] and geometrically corrected using PARGE [26].

Realistic irradiance, transmittance and atmosphere radiance samples for the PCA analysis below were obtained using a set of realistic parameters in MODTRAN-4 covering a wide range of illumination/observation angles at different times of the day and with different atmospheric conditions. In this case, given the reduced spatial extent of the DEMMIN test site, irradiance, transmittance and atmosphere radiance data were assumed to be spatially invariant since atmosphere is usually assumed to be stable over limited areas of typically few kilometers [27]. However, note that the proposed PCA formulation below is ready to be used with more general irradiance, transmittance or atmosphere radiance data if available. For instance, it is known that water vapor and other gases have more complicated spatial patterns than other causes of the transmittance [28]. This additional complexity just would imply spatio-spectral basis functions with the appropriate spatial oscillation. The same applies for differences in transmittance across the image due to large fields of view at low altitudes.

A. Eigen-analysis of the sources

Figure 1 shows the eigenfunctions and eigenvalues of the considered samples for the irradiance, reflectance, transmittance and atmosphere radiance. As in [15], training samples are used to compute the covariance matrix for each source and the spectral or spatio-spectral bases come from the diagonalization of such matrices. In all cases, the basis vectors are oscillating functions in the spatio-spectral domain. The frequency of such oscillations increases with the eigenvalue, meaning that most of the energy of the sources is concentrated in the low spatial and spectral frequencies: the sources are spectrally and spatially smooth, which is consistent with the results in [15]. This kind of behavior has also been reported for real transmittance values in the considered spectral range coming from different aerosol concentrations [29]. The shape of the basis functions is consistent with the classical result that relates periodic eigenfunctions (similar to DCT basis functions) with smooth sources [30].

One important conclusion obtained from this analysis is that the sources are band-limited. The bandwidth of the sources, the number of relevant coefficients in the (Fourier-like) PCA domain, determines the amount of data needed to characterize the sources with certain precision. Therefore, the complexity of the problem (the number of degrees of freedom to be determined) depends on this bandwidth.

B. Complexity of the measures in the spatio-spectral and spectral-only representations

In this section we show that a different number of independent data from the sources needs to be known to reproduce the

1 ATCOR4 is based on MODTRAN-4 and performs a scan angle atmospheric correction taking into account flight altitude an illumination conditions. Water vapour is estimated directly from the image using 940 nm absorption band and aerosol type and visibility estimated by field spectroscopy and meteorological auxiliary data acquired by AGRISAR ground team.
In order to compare different processing schemes, such as pixel-based versus patch-based, one has to truncate the different PCA expansions to obtain the same reconstruction error $\varepsilon$ in the measurements (using Eqs. 3 and 5). Given a target accuracy, $\varepsilon$, the number of terms required in the truncation will tell us about the intrinsic complexity (number of unknowns) of the problem in each setting. This complexity is what will eventually determine the curvature of the likelihood, and hence the difficulty of each problem or the robustness of the solution in each setting.

Figure 2 shows illustrative examples of the errors in the measurements given the truncation of the unknowns for spectral and spatio-spectral approaches (pink and gray surfaces respectively) and different amounts of spatial information (the different plots). The intersection of the error surfaces with the required representation accuracy (yellow planes) determines the number of terms to be preserved in each case. Highlighted points (in pink and gray) refer to the number of terms in the expansions required to achieve the target accuracy.

When addressing the inverse problem considered here, the error surfaces are actually defined in a 4-dimensional domain $(d_x, d_t, d_a, d_s)$, and this is how they have been used in the experiments. Nevertheless, for the sake of clarity, in this illustration of the concept, (1) we restricted the visualization of these surfaces just to two dimensions: $(d_x, d_t)$, (2) we just show the numerical values for $d_a$, and (3) we did not truncate $\mathcal{A}$ (i.e. we took $d_a = m^2 B$ in all cases).

This visualization is enough to illustrate the general trend: it takes less coefficients to achieve the desired accuracy in the spatio-spectral representations (gray curves versus pink curves) since the spectral-only error surfaces (in pink) are always over the spatio-spectral surfaces (in gray). This is consistent with the results reported in coding hyperspectral data: this behavior is the reason why it is better to jointly encode spatio-spectral patches than doing a pixel-wise encoding [12]–[15]. The particular surfaces in Fig. 2 are the average over 20 patches of size $m \times m$ (or the corresponding pixels in the spectral-only case), and were obtained using a particular choice for the spectral sensitivities (Fig. 3). However, note that the general trend (less error in the spatio-spectral case and hence less degrees of freedom in each block) is fairly independent of the selected spectral sensitivities and number of bands in accordance with the coding results [12]–[15], and it also holds when truncating $\mathcal{A}$.

From this illustration of the degrees of freedom one may expect that the retrieval problem will be better conditioned in the spatio-spectral approach, as shown in Section V both theoretically (likelihood curvatures) and in practice (actual reconstructions).

V. EXPERIMENTS

In Section III we presented the log-likelihood to be analyzed in order to assess how well conditioned the retrieval problem is. As we said, the robustness of the solution should depend on the the complexity of the sources to be recovered in the
selected (spatio-spectral or spectral-only) representation. In Section IV we saw that, in agreement with coding results, the complexity of the sources in the spatio-spectral representation is smaller than in the spectral-only case. In this section we present different retrieval problems to illustrate the positive effect of considering spatial information. The problem difficulty is first theoretically assessed according to the presented formulation (curvature of the log-likelihood) in the spectral-only and the spatio-spectral cases. Afterwards, equivalent numerical techniques are applied in each case to actually obtain the solutions confirming the expectations anticipated in the theoretical analysis.

A. Theoretical robustness of spectral-only and spatio-spectral solutions

In the examples for this illustration we assume the set of spectral sensitivities in Fig. 3. We compute the curvature of the problem as we increase the amount of considered spatial information: from the single-pixel, spectral-only, case to spatio-spectral cases with \( \alpha \in \{2, 3, 5\} \). As in Section IV we have samples from the separate elements of the problem, \((E, T, A, S)\). This is equivalent to having field measurements from a campaign before the actual mission. We assess the difficulty of the problem using such training samples through the analysis of the Hessian (computed according to Appendix A) at a number of actual solutions \((e, t, a, s)\).

2In addition to the retrieval results presented in this experimental section, supplementary material is available on-line [23]. This supplementary material consists of results including (i) retrieval over sites with substantially different spatial structure, (ii) different spatio-spectral resolution, (iii) different wavelength range. These results are consistent with the presented here since the proposed analysis is independent of issues (i)-(iii).

![Fig. 2. Degrees of freedom of the problem revealed by the truncation error, \( \varepsilon \), as a function of the number of coefficients \( d_e, d_s, d_t \) considered in the expansion of \( E, S \) and \( T \) in the spatio-spectral case (in gray) and the spectral-only case (in pink) for different block sizes: \( m = 2 \) (top-left), \( m = 3 \) (top right), \( m = 5 \) (bottom-left) and \( m = 7 \) (bottom-right). The truncation error is a scalar field in the 3D space defined by \( d_e, d_s \) and \( d_t \). For visualization purposes, in these examples we fix some truncation value for the illuminant \( d_e \), we keep all the coefficients for the atmosphere radiance \((d_s = m^2 \cdot B)\), and we show the error surfaces for the other parameters, \( d_s \) and \( d_t \). In the pixel-wise case, the number of required terms is given by the number of unknowns for one pixel times the number of pixels in the patch. This is why, for instance, taking 5 coefficients for the illuminant implies \( d_e^{\text{spatio-spectral}} = 5 \) and \( d_e^{\text{spectral}} = 45 \).]

![Fig. 3. Spectral sensitivities used in the experiments, \( R_e(\lambda) \), tuned to different bands for every pixel of the sensor array.]

![Fig. 4. shows the Hessian matrices computed for a representative patch in the 3 × 3 spatio-spectral case, and for the corresponding set of 9 pixels in the spectral-only case. In each case, for a fixed target accuracy, we obtained the number of necessary coefficients in the PCA expansions of \( E, T, A \) and \( S \) from the intersection with the 4-dimensional rate distortion surfaces, as in the illustration of Fig. 2. In this Hessian example, which is representative of the general behavior across our training scenes, it is clear that while in the spatio-spectral case the relations between the reflectance at different spatial positions are taken into account (non-diagonal matrix corresponding to coefficients of spatially extended basis functions), in the spectral-only case blocks out of the diagonal are zero since no spatial relations are taken into account. The same general behavior is obtained when considering bigger patches.]

![Fig. 5. shows the logarithm of the Hessian eigenvalues (curvatures of the likelihood) for 140 patches of different sizes. The spectral-only case and the spatio-spectral case are shown in blue and red respectively. In these plots bad conditioning behavior across our training scenes, it is clear that while in the spatio-spectral case the relations between the reflectance at different spatial positions are taken into account (non-diagonal matrix corresponding to coefficients of spatially extended basis functions), in the spectral-only case blocks out of the diagonal are zero since no spatial relations are taken into account. The same general behavior is obtained when considering bigger patches.]

![Fig. 6. shows the degrees of freedom of the problem revealed by the truncation error, \( \varepsilon \), as a function of the number of coefficients \( d_e, d_s, d_t \) considered in the expansion of \( E, S \) and \( T \) in the spatio-spectral case (in gray) and the spectral-only case (in pink) for different block sizes: \( m = 2 \) (top-left), \( m = 3 \) (top right), \( m = 5 \) (bottom-left) and \( m = 7 \) (bottom-right). The truncation error is a scalar field in the 3D space defined by \( d_e, d_s \) and \( d_t \). For visualization purposes, in these examples we fix some truncation value for the illuminant \( d_e \), we keep all the coefficients for the atmospheric radiance \((d_s = m^2 \cdot B)\), and we show the error surfaces for the other parameters, \( d_s \) and \( d_t \). In the pixel-wise case, the number of required terms is given by the number of unknowns for one pixel times the number of pixels in the patch. This is why, for instance, taking 5 coefficients for the illuminant implies \( d_e^{\text{spatio-spectral}} = 5 \) and \( d_e^{\text{spectral}} = 45 \).]
curvature in the likelihood), i.e. by the values at the left of each spectrum.

Nevertheless, the important feature of these spectra is not the absolute scale [16], but the difference between the smallest eigenvalues and the next (the smallest curvatures). While in the spatio-spectral approach just one or two of the eigenvalues are very small (there is a substantial increase in the second or third eigenvalue), in the spectral-only approach a number of them share very small values.

Zero (or close-to-zero) eigenvalues imply no curvature in the associated directions (associated eigenvectors of the Hessian). In that case, the solution is strictly undefined since there is a whole family of possible solutions (along the ridge of maximum values in the posterior). However, having only one or two zero (or very small) eigenvalues is not a problem. It is just an indetermination in the relative scale of the spectral sources we are considering: for instance, two solutions for $S \cdot E$ and $T$ give rise to exactly the same measurements if they differ on the relative scale of $S \cdot E$ and $T$. The same indetermination can take place between the scales of $S$ and $E$. However, since it is the spectral distribution what actually matters, these unavoidable sources of indetermination are not a major problem. Therefore, the real problem is having more than two zero (or very small) curvature directions. That is consistently the case in the spectral-only approach. This implies that the pixel-wise approach intrinsically leads to an ill-conditioned problem, so one can anticipate more unreliable solutions.

Interestingly, taking into account more spatial information increases the difference between the lower eigenvalues and the next, i.e. it increases the curvature of the likelihood (at least till $m = 5$, since it seems to decrease at $m = 7$ - we will come back to this point later-). This evolution of the curvature as a function of the patch size proofs that additional spatial information (up to some limit) leads to a better conditioned problem thus solving the irradiance-reflectance-transmittance ambiguity.

### B. Practical performance of the approaches

In practice, the maximum likelihood solution has to be searched in the parameter space. The theoretical results in the previous section, Fig. 5, which is consistent with the complexity example of Section [LV] indicate that the retrieval problem is better conditioned if a spatio-spectral approach is taken. This suggests that no matter the search procedure, the solutions will be more reliable and robust to noise in spatio-spectral scenarios. Here we explicitly show that this is the case in practice thus confirming the accuracy of the theoretical predictions.

In order to illustrate this point we will take the most rudimentary and reliable search technique: the exhaustive search on a discrete grid. In this setting the sources of error are: (1) the discrete nature of the search space (equivalent to the stopping tolerance condition in more sophisticated search procedures), and (2) the noise in the measurements. Of course more practical or sophisticated search techniques such as gradient descent or simulated annealing could be used. On the other hand, other completely different approaches (e.g. estimation using Markov models using different contextual neighborhoods) are possible to show the benefits of spatial information. Nevertheless, since the purpose of this section is checking the results of the curvature analysis, search in the parameter space is the most direct way to stress the point. And more sophisticated search techniques will also suffer from tolerance-like and noise issues.

1) **Retrieval procedure:** The performed search starts by randomly initializing the reflectance, the transmittance, and atmosphere radiance, $S_0$, $T_0$, and $A_0$. In each case, each random sample is independently drawn from the corresponding marginal PDFs (available from the training stage used to estimate the PCAs). See the supplementary material [31] for examples of these marginal PDFs, which are heavy tailed distributions in agreement to previous results [15].

With this initialization we exhaustively search the coefficients of $E$ on a discrete grid to minimize the log-likelihood (maximize the likelihood). We choose to search the $E$ subspace first because of its lower dimensionality. From all the equivalent configurations of number of coefficients with equivalent error (e.g. pink and gray lines in Fig. 2) we choose the lower dimensionality cases (e.g. highlighted pink and gray dots in Fig. 2). Discrete values for the considered coefficients were uniformly distributed according to the variance ranges. Note that the variance ranges are known from the PCA analysis carried out from the training samples in Section [LV]. The same grid resolution was used in the spatio-spectral and the spectral-only cases. From this estimation, $E_1$, and the randomly chosen transmittance, $T_0$, and atmosphere radiance, $A_0$, we computed $\hat{S}_1^T$ using the pseudoinverse of the sensitivities $R$. Comparison between $m = 1$ and $m > 1$ is fair since we checked that the error introduced by the pseudoinverse was the equally negligible in both the spatio-spectral and the spectral-only cases.

The estimations, $E_1$, and $\hat{S}_1^T$, together with the previous initialization $A_0$, are used in the new exhaustive search over the $T$ coefficients space, so that the likelihood is optimized again. As in the irradiance case, the number of coefficients (the dimension of the search space) is also determined by the...
The main conclusion is that, as anticipated by the proposed theoretical analysis of the likelihood curvature, considering spatial information gives rise to better retrievals over a wide range of noise levels, i.e. improved robustness. Consideration of spatial information is essential to largely improve the results. This is specially apparent in $S$ and $E$. In the case of the atmosphere transmittance and radiance the RMSE of the different approaches only differs when performance is compromised by noise in the signal (favoring spatio-spectral approaches). Additional differences between the approaches are more clear when considering not only the average RMSE, but also the variance and the nature of the errors in the individual estimations. Examples in Figs. 8, 10 illustrate this point.

Figs. 8, 10 show specific examples of the retrieved $S$, $E$, $T$ and $\mathcal{A}$ for different noise levels using the spectral-only and different spatio-spectral approaches with increasing spatial information. These examples include PSNR=$\infty$, i.e. ideal noise-free case (Fig. 8), PSNR=60 dB (Fig. 9), and PSNR=40 dB (Fig. 10). Retrieved reflectance values have to be compared to the ground truth image in Fig. 7.

In general, the inclusion of spatial information always leads to better retrieval results, specially in $S$ and $E$. A general trend of the spectral-only approach is that pixel-wise differences in
the $E$ and $T$ estimates (see the large deviation bars) give rise to pixel-wise artifacts in the spatial distribution of the retrieved reflectance even for moderate noise levels. Even though the spatial structures are better retrieved in the $m = 2$ case than in the spectral-only case, results are far from satisfactory: retrieval of $S$ is poor and the variance of the $E$ estimates is very large (i.e. large differences between the estimations in different blocks). As a result in noisy situations (e.g. Figs. 9 and 10) block-wise artifacts are clearly visible with $m = 2$.

When bigger patches are considered, more accurate and consistent estimations for $E$, $T$ and $A$ are obtained (see the reduction of the deviation bars when increasing the spatial information), leading to better retrieval of the spatial structures in $S$ and alleviating the block artifacts. As anticipated by the difference between curvatures in the $5 \times 5$ and $7 \times 7$ cases (plots at the right in Fig. 3), it seems the better conditioning is obtained for $5 \times 5$ patches, i.e. there is a plateau in the improvement due to bigger patches. Consistently, the effect of noise is more severe in the $7 \times 7$ situation (see the errors in the reflectance in Fig. 10) and see how the $7 \times 7$ error increase for lower PSNR in Fig. 9.

C. Discussion

As suggested by the retrieval results, while spatial size matters, (beyond certain limit) bigger may not be practical. It is true that the number of unknowns per pixel decreases with block size $m$, however, the dimension of the search space in the estimation problem also increases with $m$. For example, if a $15 \times 15$ hyperspectral measurements cube has to be addressed, it can be done by solving 9 separate estimation problems in $5 \times 5$ patches or a single estimation in the $15 \times 15$ patch. The $5 \times 5$ approach may require to solve 10 unknowns per patch (a total of 90 unknowns in the cube) while the $15 \times 15$ approach may only require to solve 40 unknowns. However, note that each separate estimation problem is 10-dimensional in the $5 \times 5$ case, while the single estimation problem in the $15 \times 15$ case is 40-dimensional. Accordingly, the separate $5 \times 5$ problems may be more computationally convenient than the single $15 \times 15$ problem. Therefore, the eventual gain due to the bigger size may not be worth pursuing because of the increased computational cost or the errors associated to search procedures in high-dimensional spaces. This computational issue may make a difference in practice thus inducing a (practical) upper limit to the patch size. This limit will certainly depend on the particular search procedure used in the retrieval, but it is an open issue for further research.

The goal here is just pointing out that going from the spectral-only approach to (moderate-size) spatio-spectral approaches is both theoretically convenient and feasible in practice (as illustrated even by straightforward exhaustive search).

Beyond this plateau effect, the fundamental limitation of poorly conditioned approaches (e.g. $m = 1, 2$ in the above examples) implies the need of heuristic spatial smoothness constraints to improve the results. In contrast, the consideration of additional spatial information gives rise to better results without extra spatial smoothness constraints. The proposed analysis suggests that the smoothness constraint is naturally achieved from the smooth nature of the spatio-spectral PCA basis functions.

A remark on the need of the proposed analysis: in this work we show that the ideas in [16] on the use of spatio-spectral regularities for color constancy in the context of conventional photography can be generalized to the remote sensing case. This extension is not trivial, but worth to be explicitly done, for two reasons:

- In the proposed formulation, we explicitly include the atmospheric effects which were not considered in [16] since transmittance and radiance of the media between surfaces and sensors is not an issue in conventional photography. In contrast, atmosphere transmittance is the key in remote sensing problems such weather or pollution monitoring.
- In principle the nature of the problem is different so it has to be explicitly shown that remote sensing scenes statistically behave in the same way as conventional photographic scenes. Previous results on this similarity, e.g. [15], suggest that this is going to be the case. However, note that here a specific separate analysis of light, surfaces and atmosphere, as the one presented in Section IV was strictly required.

The conclusion from the results (spatio-spectral manifolds of hyperspectral signals have reduced complexity with regard to spectral-only manifolds) is not a global criticism against works exploiting the structure of spectral manifolds in remote sensing (e.g. [32]–[34]). Actually, we see it as a positive complement of those since generic manifold learning techniques will eventually be more powerful when applied to the (simpler) spatio-spectral manifolds.

VI. CONCLUSIONS AND FURTHER WORK

The joint consideration of spatial and spectral features is crucial to solve the challenging problem of recovering the information about illumination, surface and atmosphere, from the sensor measurements. The proposed theoretical analysis of the likelihood shows that the consideration of spatial information implies better conditioning of the problem (e.g. Fig. 5). Illustrative results using equivalent numerical techniques in practical retrieval, confirm the theoretical predictions: better results are obtained in the spatio-spectral setting both in the noise-free limit and in noisy scenarios (Figs. 6, 10).

The formulation in the PCA domain implies that the analysis can be easily interpreted in statistical terms. As pointed out in Section IV the bigger explanatory power of spatio-spectral basis functions with regard to the spectral-only basis functions (lower truncation error in scenarios such as Fig. 2) implies that the actual number of unknown coefficients (effective unknowns) is smaller in spatio-spectral representations. The bigger the energy compaction in the transform, the less number of unknowns. As a result, better transforms in coding terms imply better representations for the estimation of light, surface and atmosphere from the same set of measurements. The proposed PCA analysis, namely the shape of the basis functions and the trend of the eigenvalues of the sources (Fig. 1), provides additional insight into the problem: consideration
of a limited number of coefficients is equivalent to enforcing smoothness since larger eigenvalues correspond to (spatial and spectral) low frequency patterns. Therefore, the spatio-spectral approach prevents the need of extra regularization terms heuristically enforcing spatial smoothness since these constraints naturally appear through the spatially extended basis functions. This intuition is confirmed by the smooth estimations obtained from large patches in Figs. 8-10.

Finally, as suggested in the final paragraph of Section III, since the likelihood and Hessian not only depend on the patch size, but also on other design parameters (number and spectral sensitivity of the sensors), the proposed analysis opens new possibilities for further work in sensor design: the proposed analysis can be used to determine the minimum number of sensors (and their spectral sensitivity) and the minimum size of the considered spatial region to appropriately constraint the solution of the inverse problem for a given scene variability (described by the eigenvectors of the sources). For example, while for a small set of broad-band spectral sensors the solution may not unique in spectral-only approaches, it may be well conditioned by using the appropriate spatio-spectral representation. Alternatively, if not enough spatial information is taken into account a bigger number of narrow band sensors may be required to properly constraint the solution.

APPENDIX

Here we show the generalized version of the curvature of the log-likelihood presented in [16] by explicitly including the

Fig. 8. Retrieval results in noise-free limit (PSNR = ∞ dB in the measurements). Left column, S. Second column, E. Third column, T. Right column A. The different rows correspond to progressive increase of considered spatial information. Top, spectral-only approach (no-spatial information, i.e. m = 1). Following rows: m = 2, 3, 5, 7.
atmosphere radiance and transmittance of interest in remote sensing:

\[ H_{\lambda} = \frac{\partial^2 L}{\partial e^\lambda \partial e^\mu} = 0 \]

\[ H_{\lambda} = \frac{\partial^2 L}{\partial e^\lambda \partial e^\mu} = 2 \sum_b \left( \sum_{j,k} G_{t,jk,b} \lambda^2 s^e_k \right) \left( \sum_{j,k} G_{t,jk,b} \lambda^2 s^e_k \right) \]

\[ H_{\lambda} = \frac{\partial^2 L}{\partial t^\lambda \partial t^\mu} = 2 \sum_b \left( \sum_{i,k} G_{t,im,b} \lambda^2 s^e_i \right) \left( \sum_{i,k} G_{t,im,b} \lambda^2 s^e_i \right) \]

Fig. 9. Retrieval results in low noise (PSNR = 60 dB in the measurements). Left column, S. Second column, E. Third column, T. Right column, A. The different rows correspond to progressive increase of considered spatial information. Top, spectral-only approach (no-spatial information, i.e. \( m = 1 \)). Following rows: \( m = 2, 3, 5, 7 \).
Fig. 10. Retrieval results in moderate noise (PSNR = 40 dB in the measurements). Left column, S. Second column, E. Third column, T. Right column \( \mathcal{A} \). The different rows correspond to progressive increase of considered spatial information. Top, spectral-only approach (no-spatial information, i.e. \( m = 1 \)). Following rows: \( m = 2, 3, 5, 7 \).
\[ H_{q^{(a)}}^{s_{(a)}} = \frac{\partial^2 L}{\partial q^{(a)} \partial s_{(a)}} = 2 \sum_b \left( \left( \sum_{i,k} G_{ik,b} e^{q^{(b)} q^{(a)}}(A_{ik}) \right) \right) \]

\[ H_{q^{(a)}}^{s_{(b)}} = \frac{\partial^2 L}{\partial q^{(a)} \partial s_{(b)}} = 0 \]  

\[ H_{q^{(b)}}^{s_{(b)}} = \frac{\partial^2 L}{\partial q^{(b)} \partial s_{(b)}} = 0 \]

\[ H_{q^{(b)}}^{s_{(a)}} = \frac{\partial^2 L}{\partial q^{(b)} \partial s_{(a)}} = 2 \sum_b \left( \left( \sum_{i,j} G_{ijm,b} e^{q^{(b)} q^{(a)}}(A_{ij}) \right) \right) \]

As stated above, the location superscripts, \( p, q \), are not strictly required: in the spatio-spectral case the coefficients are the same for every location in the patch (so it makes no sense to consider crossed derivatives between different locations), and in the spectral only case, no interaction between different locations is considered (as expected in a pixel-based approach with iid Gaussian noise), so the crossed derivatives are zero, and the sums over space have only one term, the one for the considered pixel. This gives rise to block diagonal Hessian matrices in the spectral-only case (see Fig. [4]).

ACKNOWLEDGEMENTS

The authors wish to thank Prof. José Moreno for his insightful comments on the atmosphere effects.

REFERENCES