



**HAL**  
open science

# Hyperspectral Texture Metrology Based on Joint Probability of Spectral and Spatial Distribution

Rui Jian Chu, Noël Richard, Hermine Chatoux, Christine Fernandez-Maloigne, Jon Yngve Hardeberg

► **To cite this version:**

Rui Jian Chu, Noël Richard, Hermine Chatoux, Christine Fernandez-Maloigne, Jon Yngve Hardeberg. Hyperspectral Texture Metrology Based on Joint Probability of Spectral and Spatial Distribution. IEEE Transactions on Image Processing, 2021, pp.4341-4356. 10.1109/TIP.2021.3071557. hal-03463722

**HAL Id: hal-03463722**

**<https://hal.science/hal-03463722>**

Submitted on 2 Dec 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Hyperspectral Texture Metrology Based on Joint Probability of Spectral and Spatial Distribution

Rui Jian Chu, Noël Richard, Hermine Chatoux, Christine Fernandez-Maloigne, Jon Yngve Hardeberg

**Abstract**—Texture characterization from the metrological point of view is addressed in order to establish a physically relevant and directly interpretable feature. In this regard, a generic formulation is proposed to simultaneously capture the spectral and spatial complexity in hyperspectral images. The feature, named *relative spectral difference occurrence matrix* (RSDOM) is thus constructed in a multireference, multidirectional, and multiscale context. As validation, its performance is assessed in three versatile tasks. In texture classification on *HyTextiLa*, content-based image retrieval (CBIR) on *ICONES-HSI*, and land cover classification on *Salinas*, RSDOM registers 98.5% accuracy, 80.3% precision (for the top 10 retrieved images), and 96.0% accuracy (after post-processing) respectively, outcompeting GLCM, Gabor filter, LBP, SVM, CCF, CNN, and GCN. Analysis shows the advantage of RSDOM in terms of feature size (a mere 126, 30, and 20 scalars using GMM in order of the three tasks) as well as metrological validity in texture representation regardless of the spectral range, resolution, and number of bands.

**Index Terms**—Hyperspectral imaging, texture, metrology

## I. INTRODUCTION

**H**YPERSPECTRAL imaging (HSI) is closely related to the measurement of physical surface properties and material composition. Thanks to the dense spectral sampling up to hundreds of bands, HSI is able to provide rich information that allows rapid and nondestructive assessment. However, highly accurate acquisition does not warrant highly accurate results. The full exploitation of HSI necessitates proper data interpretation and enforcement of metrological processing.

The interest of this work is surface nonuniformity or texture metrology. Under metrology, measurements are to be solely dependent on measurand and dissociated from sensor. Correspondingly, the similarity between textures is to be preserved regardless of the imaging sensor (grayscale, color, spectral etc.) with discrepancy only in accuracy. Feature interpretability is also of utmost importance for the establishment of a direct relationship between the measurement (feature) and the measurand (texture). In contrast to data-driven approach, metrological solutions allow the quantification of error, uncertainty, and bias for traceability [1]. However, many texture analysis methods are application oriented and far from metrology, thus mostly ad hoc. Even if they do start with meaningful

This work is supported by the French national projects ANR DigiPi and ERDF NUMERIC / e-Patrimoine.

R. J. Chu, N. Richard, H. Chatoux and C. Fernandez-Maloigne are with University of Poitiers, 86000 Poitiers, France (e-mail: {rui.jian.chu01, noel.richard, christine.fernandez}@univ-poitiers.fr).

H. Chatoux is with University of Burgundy, 21078 Dijon, France (email: hermine.chatoux@u-bourgogne.fr)

J. Y. Hardeberg is with Norwegian University of Science and Technology, 2815 Gjøvik, Norway (e-mail: jon.hardeberg@ntnu.no).

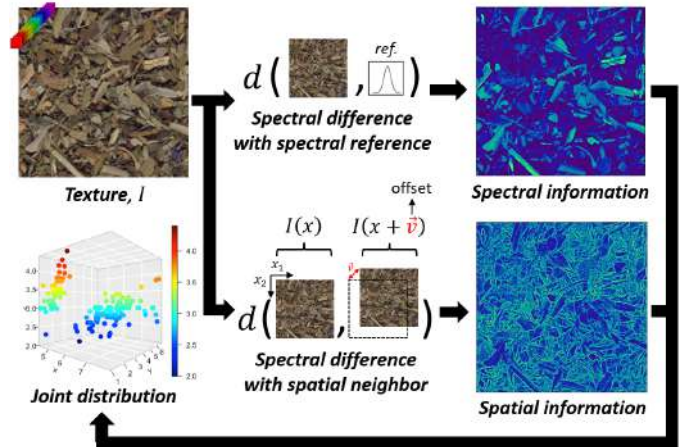


Fig. 1: Texture is defined as joint probability of spectral and spatial distribution. The feature, RSDOM consists of pixel differences with spectral reference and spatial neighbor(s).

interpretation of texture, they are often limited by empirical construction of similarity measures or complexity reduction.

The present work proposes a new perspective in hyperspectral texture analysis in the context of full-band and metrological processing which can be summarized in Fig. 1. Part of the work has been published in [2], [3], thus the main interest of this paper lies in the theoretical generalization of the feature towards multireference, multidirectional, and multiscale assessment. The contributions of this work are:

- A *generic* texture definition, encompassing a joint characterization of the spectral-spatial complexity in image.
- An multiscale and multidirectional texture feature extraction based on pixel differences, thus applicable to images of *any* spectral range, resolution, and number of bands.
- A metrological feature with a *direct* relationship with the texture's physical content for high feature interpretability.
- A rigorous assessment of the feature performance via three different tasks i.e. texture classification, content-based image retrieval, and land cover classification.

The remainder of this work is organized as follows. In Section II, a review and critical examination on state-of-the-art, which is compactly summarized in Fig. 2, is presented. A texture definition is then proposed in Section III, followed by a mathematical formulation for the feature in Section IV and V. The feature similarity measure is devised in Section VI. Section VII is dedicated to feature validation as well as interpretation. Experiments and analysis are developed in Section VIII. Finally, Section IX provides the conclusion.

## II. STATE-OF-THE-ART

### A. Tradeoff between accuracy and complexity

To take full advantage of HSI, full-band processing is necessary. Texture features defined this way [2]–[5] benefit from the high accuracy of spectral specification down to nanometres. As information is derived directly from the data, the spirit of metrology is completely preserved. The other side of the argument, however, claims that such approach leads to high computational complexity and the *curse of dimensionality* [6]–[8]. As such, dimensionality reduction is performed either by feature (band) transformation [9]–[11] or selection [12]–[14]. Using techniques e.g. principal component analysis (PCA) and independent component analysis (ICA), feature transformation projects data into a lower dimensional subspace with desirable properties such as orthogonality and independence. On the other hand, feature selection extracts a subset of the spectral bands by means of ranking, clustering, searching etc. [15].

However, the extraction of texture features preceded by dimensionality reduction is far from metrology. This is because feature transformation or selection is essentially an optimization problem subjected to the given data (data-dependent). Consequently, features extracted from different dataset are incomparable. As data is represented in the newly transformed or sampled spectral space, the underlying semantics is lost as well [16], [17]. For example, PCA has the effect of changing the geometrical properties following the scaling and centering of data. This results in the loss of physical meaning of data, prohibiting the establishment of direct relationship between texture and features as required in metrology.

### B. Multichannel consideration

Many texture extraction methods are originally developed for grayscale images. For multivariate e.g. color and spectral images, the challenge is to combine information from multiple channels. The simplest way is marginal processing i.e. by extracting feature(s) independently from each of the  $L$  channels [9]–[11]. The features are then concatenated to form a  $L$ -dimensional feature vector. Such approach is justified if the channels are independent (obtained as per PCA, ICA etc.). Otherwise, cross-channel processing is preferred as it considers the interchannel correlation [10], [12], [14]. This results in a feature vector of  $L^2-L$  dimensions as every possible pair of channels is considered. Such approach is however, statistically ad hoc with no theoretical basis in metrology.

The third approach is to consider the spectral channels as a whole or “vector-like”, which then requires the transposition of the originally two-dimensional (grayscale) methods in a three-dimensional (3-D) manner. For example, a 3-D kernel and spherical neighborhood are considered in 3-D GLCM [18] and 3-D LBP [19] respectively. For 3-D Gabor filter [20] and 3-D wavelet analysis [21], three-dimensional discrete Fourier and wavelet transform are utilized respectively.

### C. Texture feature extraction

Depending on the texture definition, features are extracted e.g. through the joint probability of pixel pairs (cooccurrence

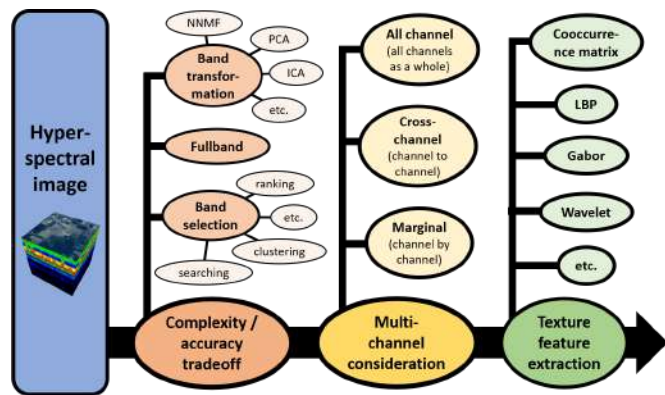


Fig. 2: Three common steps in hyperspectral texture analysis.

matrix or GLCM [22]), thresholding of neighborhood (local binary pattern or LBP [23]), spatial frequency analysis (Gabor filter [24], wavelets [25]), and model-based methods (autoregressive [26], Markovian [27], fractals [28], [29]). On the other hand, the works on procedural and dynamic textures [30], [31] are combining model-based and transform-based approaches for texture fidelity or segmentation purposes considering the correlation with human visual perception. However, the efficiency of a well-intentioned feature can be limited due to the lack of metrological considerations. For example, quantization is necessary to avoid sparsity of the GLCM for statistical reliability, thus lowering the feature accuracy. On the other hand, LBP misses out on intensity and contrast information as it considers only the sign of local differences of pixels with their neighborhood. More recently, a gravitational-based model named pattern of local gravitational force (PLGF) has been developed [32]. Though conceptually interesting, it misses out on metrology due to the empirical feature concatenation and the implicit need for quantization using histogram. Thus, it can be less adapted for multivariate e.g. hyperspectral images.

For the purpose of similarity measurement, a scalar reduction or density estimation of the feature is required. For instance, GLCM is reduced into Haralick (scalar) features [22], [33]. For transform-based approaches, moments like energy, mean, variance etc. are extracted from the filter coefficients [10], [20], [21]. The features are then concatenated to form a feature vector for which Euclidean metric is used as similarity measure [10], [20], [33]. Such practice is however, inaccurate considering the possible redundancies among the features. Moreover, the scalar reduction also reduces feature discriminability. Alternatively, the features can be modeled as histograms or probability densities [34]–[36]. Then, statistical distance e.g. divergence can be used as similarity measure. However, scale or subband independence is commonly assumed so that the densities can be modeled marginally to reduce computational complexity. This arbitrary assumption is not true in general and thus violates the spirit of metrology.

The state-of-the-art shows a lack of metrological considerations in texture analysis. To respect metrology, full-band processing is necessary with a vectorial treatment of spectral channels. A definition of “measurable” texture is then required along with a similarity measure that adheres to the mathematical nature of feature without any empirical reduction.

### III. TEXTURE DEFINITION

We begin by considering image texture as a random field, that is, a random function,  $I(x) \in \mathbb{R}^L$  over an Euclidean space,  $x \in \mathbb{R}^2$  where  $L$  is the number of spectral band. Generated by a stochastic process, the random field can be specified by its  $N^{\text{th}}$ -order joint probability distribution [37]. Commonly termed  $N^{\text{th}}$ -order statistics in the texture analysis literature [38]–[40], it refers to the probability to find at any spatial location(s) a particular  $N$ -tuple for a given spatial relationship. For example, first-order statistics is the probability to find at any spatial location a particular pixel value, whereas second-order statistics is the probability to find a particular pixel pair for a given spatial relationship. The higher the order of statistics, the better the random field is specified.

There is no consensus on the “optimum” order of statistics for universal texture discrimination [38], [41], [42]. At this point, we note that the first-order statistics provide information about the spectral distribution of individual pixels, whereas the second- and higher order ones describe the spatial distribution or interaction of the pixels. Inspired by these, we propose to define texture as the *joint probability of the spectral and spatial distribution*. Thanks to the notion of distribution, a continuous expression of the spectral and spatial properties is allowed in the spirit of metrology. In retrospect, discretization is necessary in the case of GLCM (quantization of graylevels) and LBP (encoding of local differences) which reduces the efficiency of texture characterization. Besides, a natural usage of statistical distance e.g. divergence is also possible without the need for any empirically constructed similarity measure.

In this work, we use first- and second-order statistics to quantify the spectral and spatial distributions respectively. We note that the texture definition is generic as it is applicable for image of any number of spectral band including grayscale ( $L = 1$ ) and color ( $L = 3$ ). In the following section, we develop the texture feature extraction preserving this genericity.

### IV. FEATURE EXTRACTION

#### A. Assessment of spectral distribution

The spectral distribution of an image is accessible using first-order statistics which is the probability to find at any spatial location  $x$  a particular pixel value. A direct formulation is the image histogram,  $P(I(x))$ . For multivariate image, this entails a need for multidimensional histogram. Consider a marginal quantization of  $T$  levels for each of the  $L$  bands, a formulation of  $P(I(x))$  would involve calculating and storing  $T^L$  bins. For hyperspectral image which contains up to hundreds of spectral bands, this approach is practically infeasible.

As solution, we consider a representation in the *difference* space. We represent each pixel or spectrum,  $I(x)$  by its spectral difference,  $\Delta s$  with a spectral reference,  $\hat{s}$ . Hence, the first-order statistics can be expressed as the probability to find at any spatial location  $x$  a particular spectral difference  $\Delta s$ . Hereafter denoted as  $\mathcal{J}_1$ , we express probability density function (PDF) of the continuous random variable  $\Delta s$  as:

$$\mathcal{J}_1^{\hat{s}}(I) = P\left(d(I(x), \hat{s}) = \Delta s\right). \quad (1)$$

TABLE I: NOMENCLATURE

Notation	Signification
$\mathbb{R}^D$	Set of real numbers in $D$ -dimensional space.
$I$	Image in $\mathbb{R}^L$ with $L$ bands e.g. $L = 3$ for a color image.
$x$	Spatial location in $\mathbb{R}^2$ e.g. $x = (x_1, x_2)$ .
$s$	Spectrum, considered as a continuous sampled function.
$\lambda$	Wavelength.
$\Delta s$	Spectral difference.
$\Delta G, \Delta W$	Spectral shape and intensity difference, respectively.
$\bar{s}$	Normalized spectrum, $\bar{s} = s/\omega$ .
$\hat{s}$	Spectral reference.
$\mathcal{J}$	Texture feature.
$\mathcal{J}_1, \mathcal{J}_2$	First- and second-order statistics, respectively.
$r, \theta$	Spatial distance and direction, respectively.
$d(\cdot, \cdot)$	Difference measure in $\mathbb{R}_+$ e.g. $d(I(x), I(x'))$ is a pixel difference.
$P(\cdot)$	Probability density function (PDF).
$\cap$	Logical AND operation.
$\text{KL}(\cdot\ \cdot)$	KL measure of information <sup>1</sup> , defined as $\text{KL}(f\ g) = \int f(\alpha) \log\left(\frac{f(\alpha)}{g(\alpha)}\right) d\alpha$ .
$\text{KLD}(\cdot, \cdot)$	Kullback-Leibler (KL) divergence, defined as $\text{KLD}(f, g) = \text{KL}(f\ g) + \text{KL}(g\ f)$ .
$\ \cdot\ _p$	$p$ -norm, $\ \cdot\ _p = \sqrt[p]{ \cdot ^p}$ .
$Q_z(\cdot)$	$z$ -th percentile of quantile.

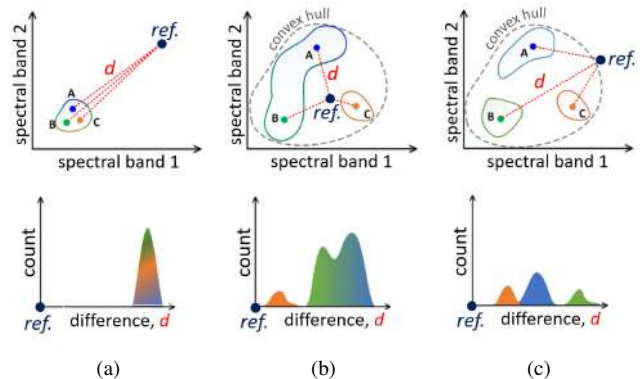


Fig. 3: An illustration on reference selection in spectral space (top row) and the resulted representation of clusters A, B, and C in difference space (bottom row). For maximum discrimination, the reference should lie just outside the convex hull of the spectral distribution as in (c). Image adapted from [44].

For maximum spectral discrimination, the spectral reference,  $\hat{s}$  has to be chosen such that it lies just outside the convex hull,  $\text{Conv}(s)$  of the spectral distribution (hereafter called *reference selection rule*) [44]. As demonstration, consider an example in Fig. 3 concerning three choices of  $\hat{s}$  for an image with two spectral bands. Suppose that the pixels can be clustered into three groups: A, B, and C. In Fig. 3 (a),  $\hat{s}$  is very far from  $\text{Conv}(s)$ . Consequently, the clusters are

<sup>1</sup>In the literature,  $\text{KL}(\cdot\|\cdot)$  is sometimes referred to as KL divergence. However, we adhere to the original work of S. Kullback and R. A. Leibler [43] which defines  $\text{KL}(\cdot\|\cdot)$  as KL measure of information and their symmetrized version as the KL divergence, the later also termed Jeffery’s divergence.

hardly distinguishable in the difference space as the spectral difference measure saturates. In Fig. 3 (b),  $\hat{s}$  is located inside  $\text{Conv}(s)$ . As a result, some of the clusters (A and B) cannot be completely discriminated as they are “equidistant” from  $\hat{s}$ . In Fig. 3 (c) whereby the reference selection rule is followed, the clusters are clearly differentiated in the difference space. Although the discrimination can be less perfect in some cases (when the spectral diversity is large) despite strict adherence to the reference selection rule, we show that this risk can be minimized using multiple references in Section V-A.

It can be shown that the point with largest aggregate difference with all other points is part of the convex hull. As such, we propose to determine the optimum spectral reference by (2). For texture discrimination across different images, a common spectral reference is required. Therefore, the calculation of (2) must be based on all the images.

$$\hat{s} = \arg \max_{s \in I} \sum_{s' \in I} d(s, s'). \quad (2)$$

### B. Assessment of spatial distribution

The spatial distribution is accessible using second-order statistics which is the probability to find a particular pixel pair for a given spatial relationship. A direct formulation is the cooccurrence matrix,  $P(I(x), I(x'))$  such that  $x$  and  $x'$  are separated at a distance  $r$  and direction  $\theta$ . However, such approach suffers the same problem as image histogram with a gigantic feature size. Furthermore, such large matrix is inevitably sparse and requires quantization for any statistical significance which in turn, accounting to information loss.

We approach such problem by again, working in the difference space. Given two random variables, their sum and difference define the principal axes of their joint probability distribution [45]. As such, cooccurrence matrix can be approximated using sum histogram,  $P(I(x) + I(x'))$  and difference histogram,  $P(I(x) - I(x'))$ . In practice, it has been shown that the difference histogram performs nearly as powerful as cooccurrence matrix in texture discrimination [46], [47]. As spectral subtraction is not physically defined, we consider instead the spectral difference  $\Delta s'$  (to differentiate from  $\Delta s$  in  $\mathcal{J}_1$ ) induced by the difference between  $I(x)$  and  $I(x')$ . Hence, the second-order statistics can be expressed as the probability to find at any paired spatial locations  $x$  and  $x'$  a particular spectral difference  $\Delta s'$ . Hereafter denoted as  $\mathcal{J}_2$ , we express PDF of the continuous random variable  $\Delta s'$  as:

$$\mathcal{J}_2^{(r, \theta)}(I) = P\left(d(I(x), I(x')) = \Delta s'\right) \quad (3)$$

s.t.  $x' = x + r \cos \theta$ .

To illustrate the relationship between spatial distribution and the PDF of pixel pair difference  $\Delta s'$ , consider the binary images in Fig. 4 consisting of just yellow and purple spectra with varying texton size. For simplicity, the spectral difference between the yellow and purple is taken as one. With large texton size (coarser texture), most pixel pairs would be of the same spectral kind. Therefore, the pixel pair difference would be mostly zero. As the texton size decreases (finer texture), more pixel pairs would be of different kind and hence the

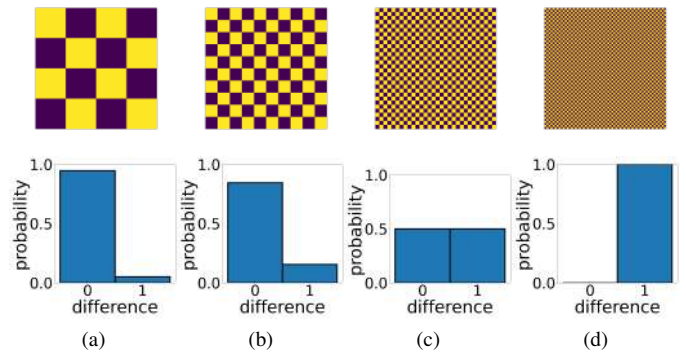


Fig. 4: Textures with varying texton size (top row) and the corresponding probabilistic distribution of pixel pair difference (bottom row), the later referred to as second-order statistics.

increasing probability of pixel pair difference being one. The same trend is observable in more complex textures too.

In an extreme case, the information captured is reduced when  $r$  is the same as texton size. In Fig. 4 (d), the spectral difference is always one as all pixel pairs are of different kind. Therefore, care has to be taken while choosing  $r$  to avoid such scenario though it can be otherwise exploited for determination of the texton size. Nevertheless, this risk can be minimized with multiscale assessment as detailed in Section V-B.

### C. Metrological calculation of pixel difference

The spectral difference<sup>1</sup> measure,  $d(\cdot, \cdot)$  is at the core of  $\mathcal{J}_1$  and  $\mathcal{J}_2$  formulation. At metrological level, the choice of  $d(\cdot, \cdot)$  cannot be disassociated with the physical definition of spectrum. To obtain a texture feature independent from the sensor resolution, we consider spectrum as a continuously function,  $s = \{s_l \mid l \in [1, L]\}$  sampled over the wavelengths,  $[\lambda_{\min}, \lambda_{\max}]$  with  $L$  spectral bands which are highly correlated.

In previous work [44], [48], it has been demonstrated that the Kullback-Leibler pseudo-divergence (KLPD) [49] is the only measure that satisfies all the metrological properties. This is because KLPD takes into account the spectral correlation which is otherwise not considered by 2-norm measures such as root mean squared error (RMSE) and spectral angle mapper (SAM) [50]. As the spectral variability of a given pixel can only be described by randomness [51], an information theoretic approach is required in place of deterministic methods i.e. RMSE and SAM. Respecting these physical constraints, KLPD is highly adapted thanks to its difference measurement based on the probabilistic discrepancy between the spectra.

KLPD is obtained by adapting Kullback-Leibler (KL) divergence [43] in the context of functions having integral unequal to one. It requires to define a normalized spectrum,  $\bar{s}$ :

$$\bar{s} = \frac{s}{\|s\|_1} \quad \text{s.t.} \quad \|s\|_1 = \int_{\lambda_{\min}}^{\lambda_{\max}} |s| d\lambda. \quad (4)$$

By choosing a continuous expression at this level, we assume that the integral will be transformed into a discrete sum taking care about the spectral sampling. As the integral of  $\bar{s}$  is equal to one, KL divergence can then be applied. Such

<sup>1</sup>The term *difference* is selected in reference to the CIE term of *color difference*. It includes in a common notion the distance and (dis)similarity.

mathematical construction allows to define KLPD as the 1-norm of the spectral shape difference,  $\Delta G(s, s')$  and the intensity difference,  $\Delta W(s, s')$  between two spectra  $s$  and  $s'$ :

$$\begin{aligned} d(s, s') &= \Delta G(s, s') + \Delta W(s, s'), \\ \text{s.t. } \Delta G(s, s') &= \|s\|_1 \cdot \text{KL}(\bar{s}\|\bar{s}') + \|s'\|_1 \cdot \text{KL}(\bar{s}'\|\bar{s}), \\ \Delta W(s, s') &= (\|s\|_1 - \|s'\|_1) \log\left(\frac{\|s\|_1}{\|s'\|_1}\right). \end{aligned} \quad (5)$$

Strictly speaking, KLPD is not a distance metric as it does not respect the triangular inequality. However, its formulation is interesting thanks to its separation of spectral difference into intensity and shape differences. To preserve this additional discriminability, hereafter we refer  $d(s, s')$  to as  $(\Delta G, \Delta W)$  i.e. a pair of differences for a bidimensional representation.

#### D. A proposed texture feature

Having defined first- and second-order statistics, we propose a texture feature named *relative spectral difference occurrence matrix* (RSDOM) using KLPD for the spectral difference measurement. In accordance to our texture definition in Section III, RSDOM conveys the joint probability to find at any spatial location  $x$  a particular spectral difference,  $\Delta s = (\Delta G, \Delta W)$  to a reference *and* a particular spectral difference,  $\Delta s' = (\Delta G', \Delta W')$  to a paired pixel located at  $x'$  for a given spatial relationship defined by distance  $r$  and direction  $\theta$ . Hereafter denoted as  $\mathcal{J}$ , we express the four-dimensional PDF of the continuous variables  $\Delta G$ ,  $\Delta W$ ,  $\Delta G'$ , and  $\Delta W'$  as:

$$\begin{aligned} \mathcal{J}^{(\hat{s}, r, \theta)}(I) &= \mathcal{J}_1^{\hat{s}}(I) \cap \mathcal{J}_2^{(r, \theta)}(I) \\ &= P \left( \begin{array}{l} d(I(x), \hat{s}) = (\Delta G, \Delta W), \\ d(I(x), I(x')) = (\Delta G', \Delta W') \end{array} \right) \end{aligned}$$

$$\text{s.t. } x' = x + r \cos \theta. \quad (6)$$

The formulation of RSDOM as a joint PDF serves three advantages. First, the expression is mathematically sound in contrast to other empirical means such as concatenation or addition of the two statistics. Second, the resulted feature is natural, intuitive and allows metrological understanding. Third, the feature is highly discriminative thanks to the dual texture representation in terms of spectral and spatial properties.

The expression of RSDOM in difference space has several desirable effects. First of all, this allows the characterization of spectral and spatial properties in a continuous manner. As the spectra are not explicitly modeled (but rather the spectral differences), quantization is unnecessary as in the case of cooccurrence matrix. Hence, information about the texture is fully preserved. Furthermore, it allows full-band processing of RSDOM without the need for dimensionality reduction e.g. PCA. Therefore, physical fidelity is preserved in line with metrology. Besides, the feature calculation is independent from the sensor resolution. This enables texture comparison across different dataset with varying spectral resolution. By extension, such difference-based assessment is also applicable to other image domains including grayscale, color, and multispectral. With adapted pixel difference measure e.g. CIELAB

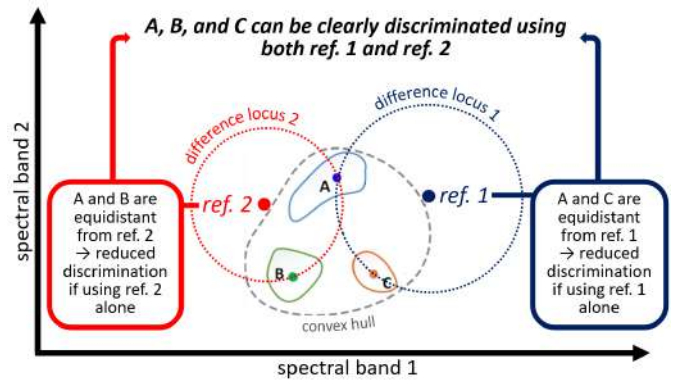


Fig. 5: An illustration on multireference spectral characterization for an image with two bands. The discriminability of each reference is limited, but combined they allow a complete discrimination of the three spectral groups. The difference loci are the sets of points that is equidistant from the reference.

$\Delta E_{ab}^*$  for the case of color images, texture characterization can be performed in exactly the same manner as RSDOM.

The given formulation of RSDOM is just one mathematical expression of the proposed texture definition. For textures with less important spectral variability, RSDOM can be simplified by considering just the average spectrum [2]. This leads to a smaller feature size, resulting in a more rapid calculation and texture similarity measurement. In cases where spectral or illuminant invariance is a desired property, first-order statistics can be left out with no effect on the method of feature parametrization and texture similarity measurement.

## V. TOWARDS A RICHER FEATURE

### A. A multireference formulation

The selection of spectral reference,  $\hat{s}$  that lies just outside the convex hull of the spectral distribution is justified in Section IV-A. However, a single reference might be insufficient as demonstrated in Fig. 5. If we consider the representation in difference space as some kind of “spectral projection”, it is clear that having multiple references provides complementary “perspectives” that improves discriminability. Akin to trilateration in the Global Positioning System (GPS), other references are able to supply additional information when not all spectral groups are discriminable from any of the references alone.

Thanks to the probabilistic construct, a multireference extension is straightforward. Considering a spectral reference set,  $\hat{S} = \{\hat{s}_1, \dots, \hat{s}_H\}$  consisting of  $H$  references, the multireference first-order statistics,  $\mathcal{J}_1 \in \mathbb{R}^{2H}$  is given by:

$$\mathcal{J}_1^{\hat{S}}(I) = P \left( \bigcap_{h=1}^H \left( d(I(x), \hat{s}_h) = (\Delta G_h, \Delta W_h) \right) \right). \quad (7)$$

### B. A multiscale and multidirectional formulation

As texture varies by scale and can be directional, a multiscale and multidirectional assessment is necessary for complete texture characterization. This can be accomplished by simply evaluating RSDOM using a set of  $r$  and  $\theta$  values.

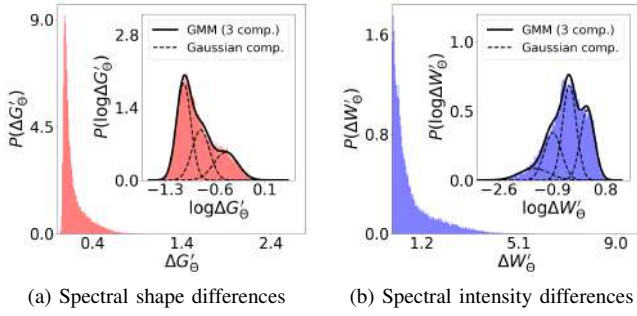


Fig. 6: A demonstration of the effect of a logarithm transformation on the distribution of spectral difference which is strictly positive. Evidently, the skewness is visibly reduced with higher compatibility for statistical modeling e.g. GMM.

Considering  $R = \{r_1, \dots, r_J\}$  and  $\Theta = \{\theta_1, \dots, \theta_K\}$ , the second-order statistics is expressed as a series of PDFs:

$$\mathcal{J}_2^{(R, \Theta)}(I) = \left\{ \mathcal{J}_2^{(j, k)}(I) \mid j \in [1, J], k \in [1, K] \right\}, \quad (8)$$

where  $\mathcal{J}_2^{(j, k)} \in \mathbb{R}^2$  denotes an evaluation on  $j^{\text{th}}$  scale and  $k^{\text{th}}$  direction. Discarding directionality information, we can consider the following simplification with rotational invariance noting that  $\mathcal{J}_2^{(r, \theta + \pi)} = \mathcal{J}_2^{(r, \theta)}$ :

$$\begin{aligned} \mathcal{J}_2^{(R)}(I) &= \left\{ \mathcal{J}_2^{(j)}(I) \mid j \in [1, J] \right\} \\ \text{s.t. } \mathcal{J}_2^{(j)}(I) &= \int_{-\pi/2}^{\pi/2} \mathcal{J}_2^{(r_j, \theta)}(I) \, d\theta. \end{aligned} \quad (9)$$

Considering a set of  $H$  references,  $\hat{S} = \{\hat{s}_h \mid h \in [1, H]\}$ ,  $J$  distances (scales),  $R = \{r_j \mid j \in [1, J]\}$ , and  $K$  directions,  $\Theta = \{\theta_k \mid k \in [1, K]\}$ , the multireference, multiscale, and multidirectional RSDOM is thus given by:

$$\begin{aligned} \mathcal{J}^{(\hat{S}, R, \Theta)}(I) &= \mathcal{J}_1^{\hat{S}}(I) \cap \mathcal{J}_2^{(R, \Theta)}(I) \\ &= \left\{ \mathcal{J}^{(\hat{S}, j, k)}(I) \mid j \in [1, J], k \in [1, K] \right\} \\ \text{s.t. } \mathcal{J}^{(\hat{S}, j, k)}(I) &= P \left( \bigcap_{h=1}^H \left( \begin{aligned} d(I(x), \hat{s}_h) &= (\Delta G_h, \Delta W_h) \\ d(I(x), I(x'_{j,k})) &= (\Delta G'_{j,k}, \Delta W'_{j,k}) \end{aligned} \right) \right) \\ \text{s.t. } x'_{j,k} &= x + r_j \cos \theta_k. \end{aligned} \quad (10)$$

## VI. SIMILARITY MEASUREMENT

Having performed feature extraction, texture is discriminated thanks to a feature similarity or distance measurement. We note that for the purpose of metrology, the similarity or distance measure must be developed respecting the mathematical nature of the feature. In this regard, the similarity measure must take into account the probabilistic nature of RSDOM.

For maximum discrimination, we decide not to impose any *a priori* assumptions but to extract information directly from the feature. Given two PDFs, their similarity can be

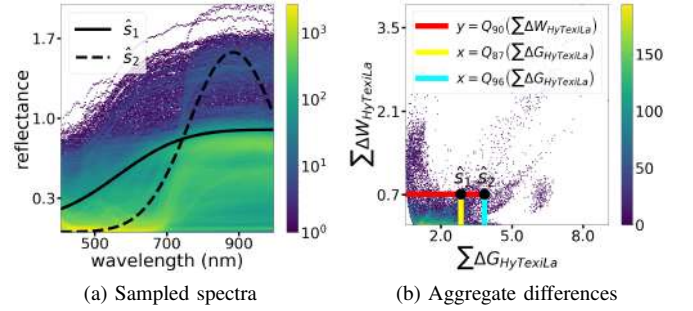


Fig. 7: The selection of spectral references,  $\hat{s}_1$  and  $\hat{s}_2$  considering 80% and 90% of the sampled spectra respectively.

measured using likelihood ratio test [52] which is a form of nonparametric statistical test. It can be shown that maximizing the log likelihood ratio is equal to minimizing KL measure of information [53]. Hence, we can use KL divergence for similarity measurement. Although it exists many other alternatives [54] e.g.  $\alpha$ -divergence and Bregman's divergence as well as  $s$ -divergence [55], the study of an optimal similarity or distance measurement is beyond the scope of this work.

Due to the multidimensionality of RSDOM, it is impractical to process the KL divergence directly on histogram features. Moreover, such approach requires the same PDF support which can vary greatly depending on the texture. As a solution, we propose to use Gaussian mixture model (GMM), a class of density estimation methods that is capable of modeling any complex PDF given sufficient number of components. Using GMM, RSDOM can be parameterized using just few parameters with a feature size of  $M(1+D+D(D+1)/2)$  where  $M$  is the number of GMM components and  $D$  is the dimensionality. As there is no closed form solution for KL divergence between GMMs, we use the variational approximation [56].

As RSDOM is a PDF of spectral differences which are positive measures by definition, it is positively skewed and dense around zero. To increase feature discriminability and compatibility to GMM, a diffeomorphism is applied by taking the logarithm of each dimension. Such transformation has a desirable effect of stretching the values close to zero while compressing those far from zero, thus reducing the skewness and improving the normality as demonstrated in Fig. 6.

## VII. FEATURE VALIDATION AND INTERPRETATION

### A. Selection of spectral reference

Our reference selection rule states that for maximum discrimination, the reference has to be chosen such that it lies just outside the convex hull of the spectral distribution. Here, we perform an illustration based on *HyTexiLa*.

*HyTexiLa* is a hyperspectral dataset of 112 images having spatial dimensions of  $1024 \times 1024$ . Obviously, it is impractical to consider all the pixels (spectra). As such, we randomly sample 250 spectra from each image as visualized in Fig. 7 (a). Next, we calculate the aggregate shape and intensity differences, hereby denoted as  $\sum \Delta G_{HyTexiLa}$  and  $\sum \Delta W_{HyTexiLa}$  respectively as shown in Fig. 7 (b) using histogram. The last step then would have been to select the spectra with the largest  $\sum \Delta G_{HyTexiLa}$  and  $\sum \Delta W_{HyTexiLa}$  as the reference.

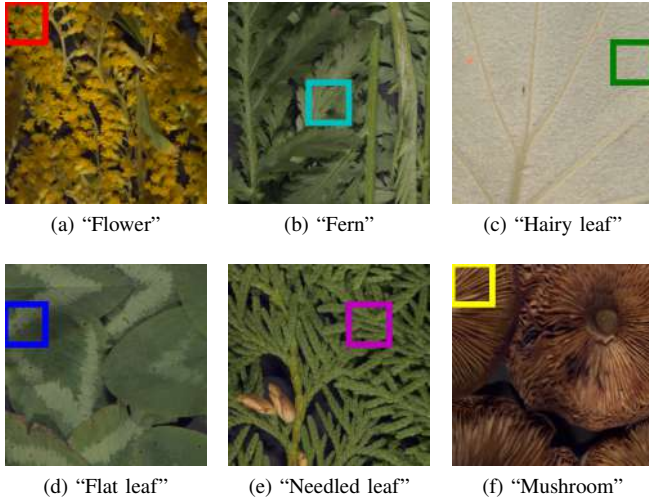


Fig. 8: Textures considered for the validation of RSDOM formulation in Table II. Bordered in color are the patches considered in the illustration of RSDOM in Fig. 9 and 10.

However, selecting reference purely in this manner would have risked choosing one that is too “far” from majority of the spectra. This is due to the large spectral diversity as in the case of *HyTexiLa* and the presence of outliers which are common in natural images. This will cause the spectral difference to saturate, hence reducing the discriminability. Therefore, we *relax* the reference selection rule by considering a subset of the spectra. As a working example, we first consider the 90th percentile of the aggregate intensity differences, hereby denoted as  $Q_{90}(\sum \Delta W_{HyTexiLa})$  as indicated in Fig. 7 (b). Next, we select the  $z$ -th (value of  $z$  to be determined) percentile of aggregate shape differences,  $Q_z(\sum \Delta G_{HyTexiLa})$  such that a defined percentage of spectra are considered (in conjunction with  $Q_{90}(\sum \Delta W_{HyTexiLa})$ ). The spectrum with  $\sum \Delta G_{HyTexiLa} \approx Q_z(\sum \Delta G_{HyTexiLa})$  and  $\sum \Delta W_{HyTexiLa} \approx Q_{90}(\sum \Delta W_{HyTexiLa})$  are then the approximate solution to the argmax problem in (2) and is taken as reference.

As the spectral characterization is performed in difference space, what matter are the spectral differences of the reference to the spectra. Therefore in practice, any spectrum can be taken as the reference as long as the spectral differences are preserved. For a multireference assessment, we choose to work with two references,  $\hat{s}_1$  and  $\hat{s}_2$  which covers approximately 80% ( $z = 87$ ) and 90% ( $z = 96$ ) of the sampled spectra respectively. For flexibility and scalability,  $\hat{s}_1$  and  $\hat{s}_2$  is defined, but not limited to the following mathematical functions:

$$\begin{aligned} \hat{s}_1 &= A_1 \operatorname{erf} \left\{ \frac{s - \mu_1}{\sigma_1} \right\} + B, \\ \hat{s}_2 &= A_2 \exp \left\{ - \left( \frac{s - \mu_2}{\sigma_2} \right)^2 \right\}, \end{aligned} \quad (11)$$

with  $A_1 = 0.4$ ,  $B = 0.5$ ,  $\mu_1 = 564.95$  nm,  $\sigma_1 = 200$  nm,  $A_2 = 1.576595$ ,  $\mu_2 = 884.12$  nm, and  $\sigma_2 = 100\sqrt{3}$  nm. The choice of  $A_2$  is to ensure that  $\int \hat{s}_1 d\lambda = \int \hat{s}_2 d\lambda$  so that  $\Delta W_1 = \Delta W_2 = \Delta W$  for all spectra. Such selection allows to illustrate the spectral distribution,  $\mathcal{J}_1$  in three dimensions for direct visualization as displayed in Fig. 9 (d).

TABLE II: VALIDATION OF RSDOM FORMULATION IN TERMS OF SPECTRAL-SPATIAL, MULTIREFERENCE, AND MULTIDIRECTIONAL ASSESSMENT AS WELL THE CHOICE OF KLPD AS SPECTRAL DIFFERENCE FOR TEXTURES IN FIG. 8.

Spectral difference	Feature	Spectral reference	Direction, $\theta$ (distance, $r = 1$ )	Accuracy
KLPD	$\mathcal{J}_1$	$\hat{s}_1$	-	$91.1 \pm 3.7$
		$\hat{s}_2$		$93.5 \pm 2.8$
		$\hat{s}_1, \hat{s}_2$		$97.6 \pm 1.6$
	$\mathcal{J}_2$	-	0	$88.0 \pm 5.6$
			$\theta = 0, \frac{\pi}{4}, \frac{\pi}{2}, \frac{3\pi}{4}$	$90.0 \pm 3.7$
	$\mathcal{J}_1, \mathcal{J}_2$	$\hat{s}_1$	0	$91.9 \pm 3.6$
			$\theta = 0, \frac{\pi}{4}, \frac{\pi}{2}, \frac{3\pi}{4}$	$95.1 \pm 2.6$
		$\hat{s}_2$	0	$94.1 \pm 4.0$
			$\theta = 0, \frac{\pi}{4}, \frac{\pi}{2}, \frac{3\pi}{4}$	$96.0 \pm 2.1$
		$\hat{s}_1, \hat{s}_2$	0	$97.6 \pm 2.0$
			$\theta = 0, \frac{\pi}{4}, \frac{\pi}{2}, \frac{3\pi}{4}$	<b><math>98.4 \pm 1.4</math></b>
	SAM	$\hat{s}_1, \hat{s}_2$	$\theta = 0, \frac{\pi}{4}, \frac{\pi}{2}, \frac{3\pi}{4}$	$96.0 \pm 2.6$
RMSE	$91.6 \pm 3.3$			
SID	$97.5 \pm 1.8$			

### B. Validation of formulation

We verify each element in the formulation of RSDOM based on six different textures: “flower”, “fern”, “flat leaf”, “hairy leaf”, “needled leaf”, and “mushroom” from *HyTexiLa* as shown in Fig. 8. The discriminability of RSDOM is assessed both qualitatively and quantitatively through illustrations of the PDFs (Fig. 9 and 10) and classifications of the textures respectively (Table II). To aid visualization, we choose to display the RSDOM only for relatively stationary part (patch) of the textures as bordered in red, cyan, green, blue, magenta, and yellow (Fig. 8). Nevertheless, all parts of the textures are considered for the classification for which the same protocol from Section VIII-A is employed. The verdicts are:

**On spectral characterization:** From Table II, it can be seen that the spectral classification using  $\mathcal{J}_1$  is quite performant (97.6%) with two spectral references. In fact, the result is close to that of spectral-spatial classification (98.4%). This demonstrates the efficiency of  $\mathcal{J}_1$  as well as spectral information being the main source of discrimination for the six textures.

**On spatial characterization:** The spatial classification using  $\mathcal{J}_2$  archives an accuracy of 90.0% with four directions assessed, lower than that of the spectral classification (97.6%). This is due to the fact texture can exhibit the same spatial variation but with different spectral characteristics. This shows that spatial assessment alone is insufficient for texture discrimination and must be aided by spectral assessment.

**On spectral-spatial characterization:** The spectral-spatial classification (98.4%) using RSDOM outperforms both spectral (97.6%) and spatial classification (90.0%). The high discriminability of RSDOM can also be observed in Fig. 9 (e) which shows maximum class separability thanks to the complete characterization in spectral and spatial variations.

**On multireference assessment:** The performance of RSDOM improves (from 91.9% and 94.1% to 97.6% for unidirectional



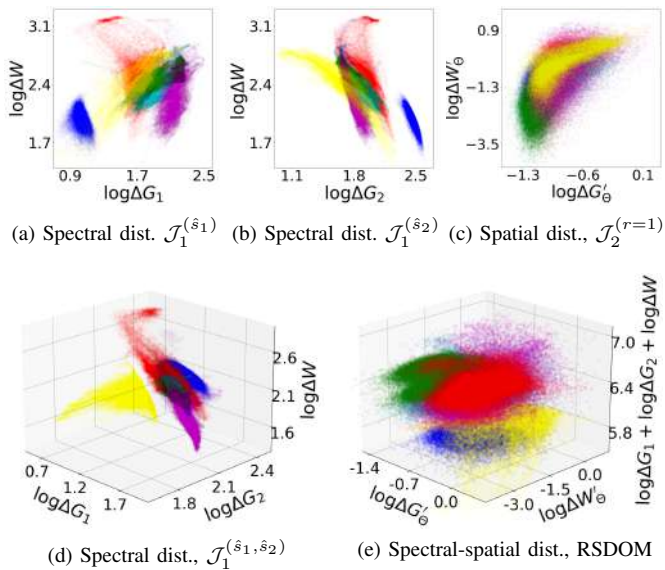


Fig. 9: An illustration of RSDOMs calculated using two spectral references,  $\hat{s}_1$  and  $\hat{s}_2$  for the patches bordered in color in Fig. 8. Different colors represent different textures. For visualization, the five-dimensional RSDOMs (joint PDF of  $\mathcal{J}_1$  and  $\mathcal{J}_2$ ) are depicted in several subspaces. The interest of the joint spectral-spatial characterization is clearly shown.

assessment; from 95.1% and 96.0% to 98.4% for multidirectional's) with two references. As demonstrated in Fig. 9 (d), the class separability increases with multiple references thanks to the additional perspectives. In particular, the yellow class (“mushroom”) is perfectly separated from others in  $\mathcal{J}_1^{(\hat{s}_1, \hat{s}_2)}$  in contrast to  $\mathcal{J}_1^{(\hat{s}_1)}$  and  $\mathcal{J}_1^{(\hat{s}_2)}$ 's. A higher locality is also observed in all classes which will aid the statistical modeling. **On multidirectional assessment:** The performance of RSDOM improves (from 91.9% and 94.1% to 95.1% and 96.0% respectively for single reference assessment; from 97.6% to 98.4% for multireference's) with multidirectional assessment. Evidently, evaluation in single direction is sufficient for isotropic textures but not for directional ones like “mushroom”. Fig. 10 (a) and (b) show the evolution of RSDOM with respect to direction,  $\theta$ . For “hairy leaf” which is isotropic, the RSDOMs are nearly identical for all  $\theta$ , demonstrating the sufficiency of a unidirectional assessment. As for “mushroom”, the RSDOMs differ according to  $\theta$  with larger values exhibited around  $\theta = 5\pi/8$ , revealing the main spatial variation in that direction. Thanks to these properties, RSDOM can also be exploited to detect and identify directionality in images.

**On multiscale assessment:** Designed for metrological purposes, RSDOM is not scale-invariant and responds positively to scale changes as required. This is demonstrated in Fig. 10 (c) and (d) which show the evolution of RSDOM with respect to distance (scale),  $r$ . Nevertheless, fractal dimension can be directly obtained from this evolution for scale invariance [57].

**On the choice of spectral difference:** It can be seen that RSDOM calculated using KLPD performs best (98.4%) thanks to its metrological decomposition of spectral difference into shape and intensity for extra discriminability. In contrast, SAM

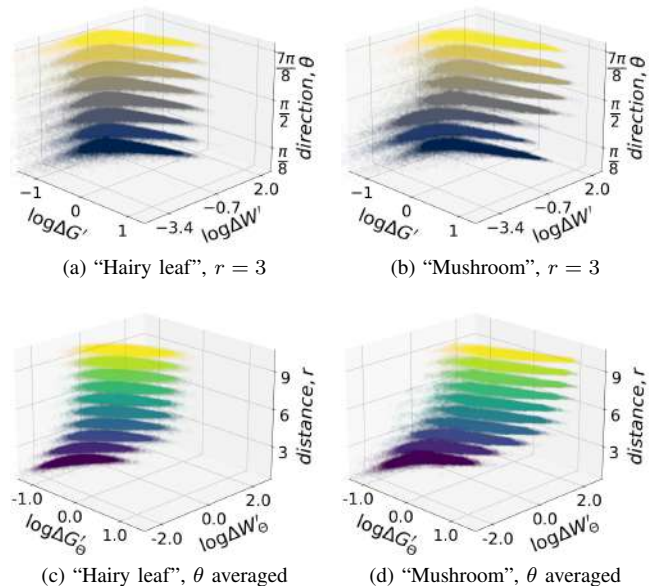


Fig. 10: Evolution of RSDOM (spectral distribution not shown) assessed with different direction,  $\theta = k\pi/8, k \in [0, 7]$  and distance,  $r \in [1, 10]$  for “hairy leaf” and “mushroom”. Designed for metrology, RSDOM responds positively to scale and direction changes. For clarity, the RSDOMs for the eight  $\theta$  values and ten  $r$  values are plotted in different colors.

[50] measures the angle between spectra by treating them like vectors. Thus by construction, SAM is invariant to the intensity information which could explain its lower performance (96.0%). RMSE or Euclidean distance, on the other hand, considers the spectral difference marginally (band-by-band). Like SAM, the spectral correlation is also ignored which possibly contributes to its lower performance (91.6%). Related to KLPD, spectral information divergence (SID) [51] also treats spectra as probabilistic constructs. It forces spectra into probabilities by dividing them with their 1-norm before directly processing their KL divergence. Consequently, spectral intensity difference is ignored. In the calculation, equal weights are also empirically assigned in the KL divergence (in contrast to the weighting by spectral intensities in KLPD, which arises naturally in a mathematical development). Lacking in theoretical coherence, it is thus less performant (97.5%).

### C. Feature interpretability

For metrological purposes, a direct relationship between the feature (measurement) and physical texture (measurand) is expected. In Fig. 11, we demonstrate the interpretation of  $\mathcal{J}_1$  and  $\mathcal{J}_2$  (of which the joint consideration constitutes RSDOM) in terms of spectral and spatial distribution. For simplicity, we consider only the spectral shape,  $\Delta G$  as the interpretation for spectral intensity,  $\Delta W$ , is similar. Firstly, the modality of  $\mathcal{J}_1$  indicates the “number” of colors in the texture. For instance, a unimodal  $\mathcal{J}_1$  indicates the presence of just one colour whereas a bimodal  $\mathcal{J}_1$  suggests a bicolor texture. Secondly, the width of a given peak in  $\mathcal{J}_1$  reflects the color variance. A narrow peak indicates a relatively “pure” color whereas a wide peak signifies the existence of multiple color

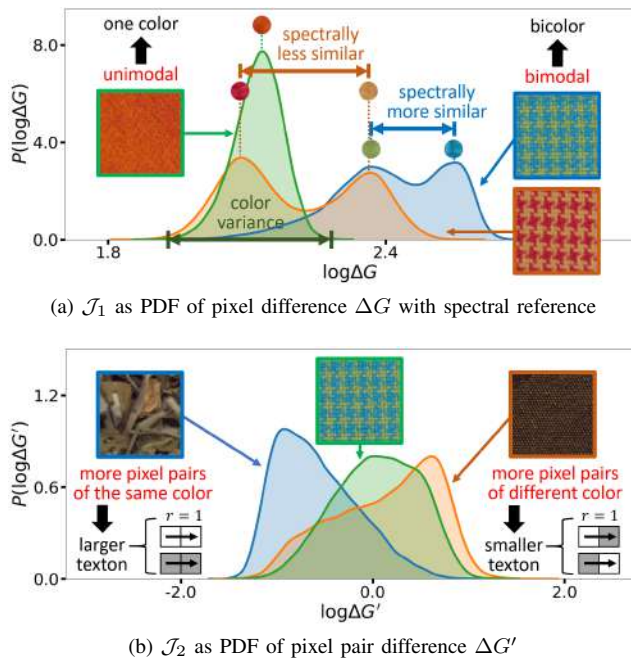


Fig. 11: A demonstration on how to interpret  $\mathcal{J}_1$  and  $\mathcal{J}_2$  (of which the joint consideration constitutes RSDOM) in terms of spectral and spatial distribution of a given texture.

tints or shades. Thirdly, the distance of a peak to the other(s) in  $\mathcal{J}_1$  illustrates their spectral similarity. The further they are from each other, the more spectrally distinct the color they represent. Fourthly, the skewness of  $\mathcal{J}_2$  conveys information on spatial interaction between the pixels for a given distance,  $r$ . A positively skewed  $\mathcal{J}_2$  is due to a higher probability of pixel pairs having the same color which indicates a relatively larger texton size. In contrast, a negatively skewed  $\mathcal{J}_2$  is due to a higher probability of pixel pairs having different colors which in turn, hints at a relatively smaller texton size.

We illustrate the interpretation of RSDOM with more examples. Fig. 12 exemplifies feature extraction using RSDOM, GLCM, Gabor filter, and LBP for three textures: “dotty” (resembling dots), “red” (red + yellow pattern), and “cyan” (cyan + green pattern). Suppose that the pixels in “dotty”, “red”, and “cyan” can be clustered into brown/white, red/yellow, and cyan/green respectively. Thanks to RSDOM, their spectral-spatial distribution can be deduced easily. For “dotty”, it can be seen that there is a very gradual brown/white transition with significantly more brown-like pixels. Besides, the large dynamic range of  $\Delta G_2$  indicates a high spectral dissimilarity for brown/white. As for “red” and “cyan”, there is a clear spectral distinction between red/yellow and cyan/green pixels as indicated by the narrow peaks in  $\mathcal{J}_1$ . However, the spectral similarity between the two colors for “red” is lower than that of “cyan” which is indeed the case for red/yellow compared to cyan/green. On the other hand, the positively skewed  $\mathcal{J}_2$  of “dotty” indicates a smaller texton size compared to “red” and “cyan” which has a nearly symmetric  $\mathcal{J}_2$ . Still,  $\mathcal{J}_2$  of “red” is slightly more positively skewed than “cyan”. This agrees well with our visual inspection that the “red” textons are somewhat larger than “cyan”’s. Clearly, RSDOM is highly interpretable with a direct relationship to the texture’s physical content.

The high interpretability of RSDOM is not shared by others in hyperspectral texture metrology. By design, GLCM summarizes the various image second-order statistical properties, Gabor energy reflects the energy or “amount” of texture for a certain scale and direction whereas LBP acts like a texton dictionary [58]. However, their interpretation is only valid within the context of a given spectral channel. It is impossible to reconcile texture information coming from different spectral channels and to interpret them. Furthermore, the features can vary greatly if different set of spectral bands is considered. As such, interpretation is also infeasible among features obtained using different spectral dimensionality reduction approach.

## VIII. EXPERIMENTS AND ANALYSIS

### A. Texture classification on HyTexiLa

In this first task, we assess the performance of RSDOM in a texture classification scheme applied on a hyperspectral dataset. As each image is treated as a unique class, the entire image is to be considered for characterization. This allows an assessment of RSDOM’s capability as a global descriptor.

**Dataset:** Composed of 112 reflectance images, *HyTexiLa* [14] is a dataset of spectrally and spatially high resolution texture from five categories i.e. food (10), stone (4), textile (65), vegetation (15), and wood (18). Each image measures  $1024 \times 1024$  pixels with  $L = 186$  spectral bands. Sampled at an interval of 3.19 nm, the wavelengths range from 405.37 nm to 995.83 nm while covering visible and near infrared regions.

For the classification, we employ a nearest neighbor (1-NN) search whereby the classification is performed based on texture similarity alone. Such nonparametric approach allows us to attribute the classification performance directly to the feature’s discriminability instead of the classifier’s efficiency. Furthermore, this also enables us to deal with the issue of nonstationarity as there is a versatility of texture representation from which each class can be identified. Each image is separated into 25 patches without overlapping of which 12 and 13 of them are randomly selected for training and testing respectively. To reduce bias, we repeat the classification 100 times and consider the average performance. Both accuracy (percentage correct classification) and  $F_1$  score (harmonic mean of precision and recall) are reported for *HyTexiLa* (intercategorical) and intracategorical classification.

**Computations:** For RSDOM, we employ a multireference, monoscale, and rotational-invariant assessment. For the spatial assessment, we choose  $r = 1$  for maximum “sensitivity” to the tiniest spatial variation assuming negligible noise level. For rotational invariance, we compute the average of spectral difference within the digital circle [59] of radius  $r$  sampled at  $K = 4$  directions,  $\Theta = \{0, \pi/4, \pi/2, 3\pi/4\}$ , hereafter denoted as  $(\Delta G'_\Theta, \Delta W'_\Theta)$ . The resulted RSDOM is a five-dimensional probability density function of  $\Delta G_1$ ,  $\Delta G_2$ ,  $\Delta W$ ,  $\Delta G'_\Theta$ , and  $\Delta W'_\Theta$ . The number of GMM components,  $M$  is selected such that the classification accuracy is maximized.

As comparison, we also consider a spectral classification (using  $\mathcal{J}_1$ ). The following state-of-the-art are too computed:

- GLCM: A spatial distance of 1 is used with 32 quantization levels. As in [18], [60], [61], five Haralick features

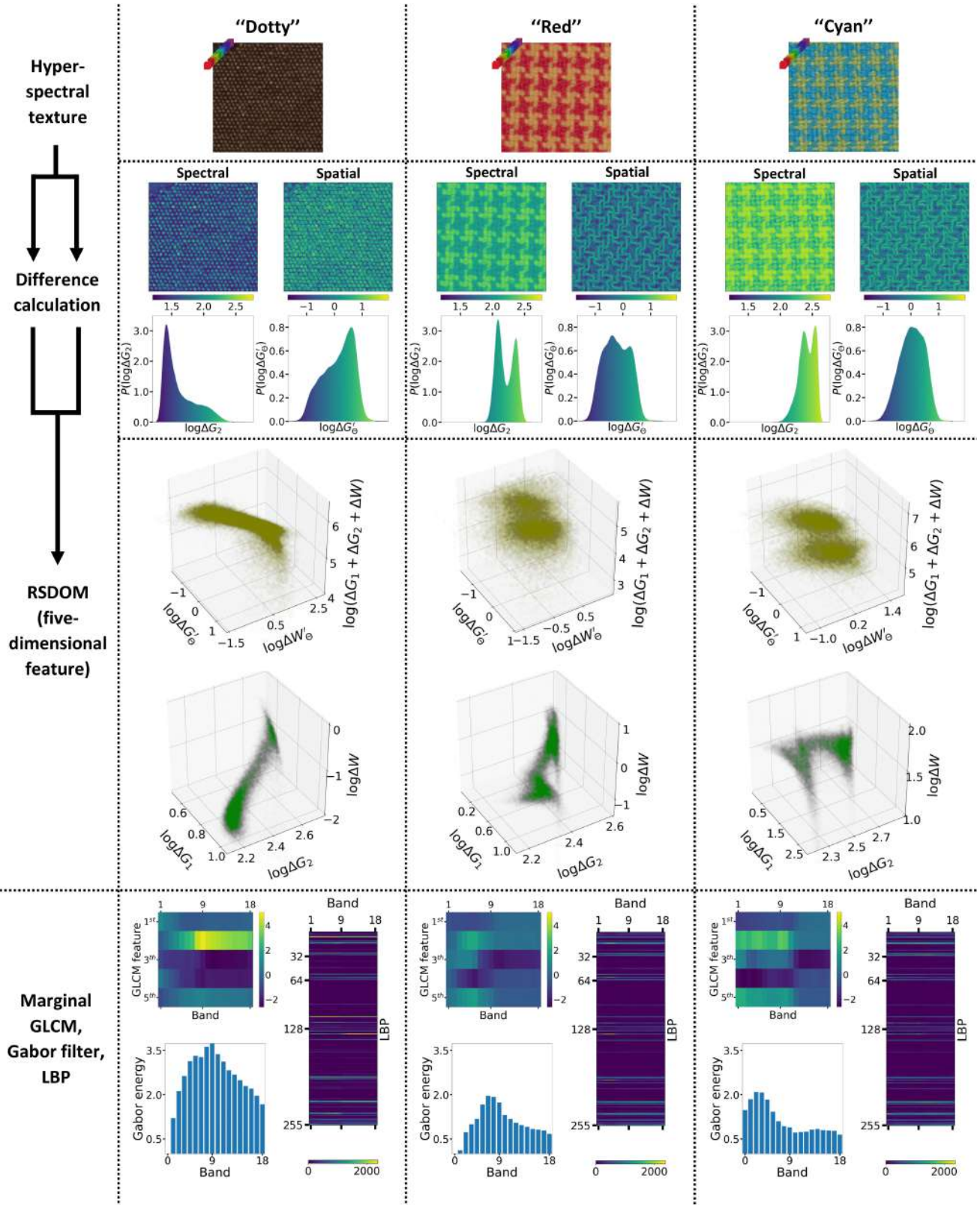


Fig. 12: Texture feature extraction using RSDOM for "dotty", "red", and "cyan" from *HyTexiLa*. **Row 1:** The hyperspectral textures. **Row 2:** The pixel difference with spectral reference (hereafter termed spectral difference) and spatial neighbors (hereafter termed spatial difference) are computed. **Row 3:** The spectral and spatial distribution can be interpreted from the probability density function (PDF) of spectral ( $\mathcal{J}_1$ ) and spatial difference ( $\mathcal{J}_2$ ) respectively. **Row 4 and 5:** RSDOM is expressed as the joint distribution of  $\mathcal{J}_1$  and  $\mathcal{J}_2$ . For visualization, the five-dimensional RSDOM is illustrated in two three-dimensional views. **Row 6:** The marginally processed GLCM, Gabor feature, and LBP are shown for comparison.

TABLE III: COMPARISONS OF ACCURACY AND  $F_1$  SCORE (ITALICIZED) FOR TEXTURE CLASSIFICATION ON *HyTexiLa*. BOLD NUMBERS INDICATE THE BEST PERFORMANCE. RESULTS AVERAGED OVER 100 CLASSIFICATIONS. INDICATED IN BRACKETS ARE THE CORRESPONDING NUMBER OF GMM COMPONENTS.

Category	No. class	<i>Full-band</i>		<i>Marginal (m-)</i>			<i>Cross-channel (cc-)</i>		
		Spectral	RSDOM	GLCM	Gabor	LBP	GLCM	Gabor	LBP
Food	10	98.0±1.1 (2)	99.7±0.5 (2)	98.3±1.1	93.2±1.8	92.5±2.4	98.2±1.2	97.6±1.4	<b>99.9±0.3</b>
Stone	4	<b>100±0.0</b> (8)	<b>100±0.0</b> (5)	87.7±3.9	91.5±3.1	94.8±2.7	95.5±2.7	91.7±3.4	99.7±0.7
Textile	65	99.8±0.2 (5)	99.9±0.2 (6)	96.3±0.6	98.1±0.7	93.0±0.8	99.2±0.3	99.2±0.4	<b>100±0.1</b>
Vegetation	15	95.5±1.6 (5)	96.9±1.3 (8)	79.3±2.5	81.5±2.8	78.9±2.8	89.3±2.1	91.2±2.5	<b>97.2±1.5</b>
Wood	18	94.5±2.4 (8)	<b>96.3±1.8</b> (6)	74.2±3.4	82.1±3.3	90.6±3.2	83.8±2.9	85.6±3.0	93.6±2.7
<i>HyTexiLa</i> (all categories)	112	97.7±0.6 (5) 97.8±0.5	<b>98.5±0.5</b> (6) <b>98.6±0.4</b>	88.0±1.0 88.4±1.0	91.6±1.2 91.8±1.1	89.1±1.3 89.4±1.3	94.7±0.7 94.9±0.7	94.7±0.9 94.9±0.8	98.4±0.6 98.5±0.5

i.e. energy, entropy, contrast, correlation, and homogeneity are considered. For feature similarity measurement, normalized Euclidean distance [33] is used.

- Gabor filter: The frequency is set at 0.5 with a bandwidth of one octave. For each spectral band, the Gabor energy is defined using 2-norm of the filter responses [10]. Normalized Euclidean distance [33] is used.
- LBP: A neighborhood of eight pixels ( $K = 8$ ) with radius 1 is considered. Histogram intersection [62] is used.

For multidirectional assessment, the GLCM features and Gabor energies are averaged over  $K = 4$  directions as RSDOM. We consider only their (and LBP) implementation in marginal (m-) and cross-channel (cc-) processing [33], [10] and [14] as their full-band formulation are metrologically impractical due to computational complexity. For dimensionality reduction, we employ band selection with uniform spectral spacing of  $\delta = 10$  bands so that the interchannel correlation is at least 0.9. The *HyTexiLa* images are thus spectrally downsampled to  $L' = 18$  bands for the computation of GLCM, Gabor filter, and LBP. In contrast, RSDOM is processed in full-band ( $L = 186$ ).

**Results:** From Table III, it can be seen that RSDOM scores excellently (98.5%) in *HyTexiLa* classification. It performs significantly better than m-GLCM (88.0%), cc-GLCM (94.7%), m-Gabor (91.6%), cc-Gabor (94.7%), and m-LBP (89.1%). Compared to cc-LBP (98.4%), RSDOM performs slightly better (+0.1%) but with a feature size 650 times smaller ( $M(1 + D + D(D + 1)/2) = 126$  scalars for RSDOM vs.  $2^K \cdot L^2 = 82944$  for cc-LBP). On computational speed, the time for calculating RSDOM for an image patch (with spatial dimensions of  $204 \times 204$  pixels and  $L = 186$  spectral bands) is measured at 3.1 s or 4.5 s with GMM ( $M = 6$ ) on an Intel i7 CPU (2.90 GHz). As for cc-LBP, the time is measured at 106.9 s for a full-band processing or 1.1 s for operation on a reduced set of  $L' = 18$  bands. Spectral dimensionality reduction is also unnecessary for RSDOM which is otherwise required for efficient implementation of cc-LBP. Full-band processed, light, and performant, RSDOM is clearly adapted for metrology.

Referring to Fig. 13, a number of factors can be attributed to the *HyTexiLa* misclassification. The main contributors are the vegetation and wood images which are highly nonstationary. For example, more than one kind of texture can present in the flower image e.g. seed's and petal's. Moreover, the black background (supposedly the imaging platform) is visible in 30%

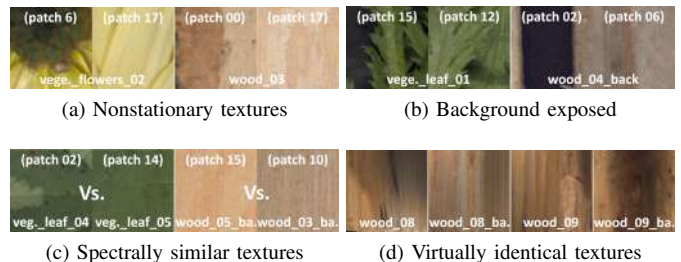


Fig. 13: Instances of *HyTexiLa* misclassification.

and 17% of vegetation and wood respectively which further promotes the nonstationarity. Besides, the vegetation and wood are predominantly green and brown respectively which poses challenges in spectral discrimination. One of the solutions is to add spectral references that lie just outside the convex hull of the green and brown spectra, although this comes at the cost of increased feature dimensionality. On the other hand, high similarities are found between “wood\_08”, “wood\_08\_back”, “wood\_09”, and “wood\_09\_back”. It is highly likely that they are taken from the same piece of wood, inducing virtually identical textures that are easily mistaken for each other as they contribute to 18.8% of the *HyTexiLa* misclassification. Last but not least, needless to say some textures would have been more efficiently discriminated with multiscale analysis. On a side note, the performance of RSDOM is subjected to the efficiency of GMM. In some cases, GMM could fail to fully capture the richness of RSDOM and thus causing its lower discriminability. An optimization on the statistical modeling is however, beyond the scope of this work.

### B. Content-based image retrieval on ICONES-HSI

In this second task, we assess the performance of RSDOM in a content-based image retrieval (CBIR) framework in which the candidate images are ranked according to their relevance to the query image. Contrary to image classification which could be optimized by using a complex classifier, image retrieval is a purely distance-based approach and is therefore highly dependent on the quality of features extracted. As the entire image is also considered, the efficiency of RSDOM as a global descriptor is again evaluated. However, its robustness aspect is emphasized as the candidate images are often very versatile.

TABLE IV: COMPARISONS OF PRECISION AT TOP 10 RETRIEVED IMAGES (P@10) AND MEAN AVERAGE PRECISION (MAP) FOR CONTENT-BASED IMAGE RETRIEVAL ON *ICONES-HSI*. BOLD NUMBERS INDICATE THE BEST PERFORMANCE. INDICATED IN BRACKETS ARE THE CORRESPONDING NUMBER OF GMM COMPONENTS.

Class	No. img.	<i>P@10</i>					<i>MAP</i>				
		Spectral	RSDOM	GLCM	Gabor	LBP	Spectral	RSDOM	GLCM	Gabor	LBP
Agriculture	50	65.6 (7)	<b>79.0 (2)</b>	58.6	24.2	43.8	40.4 (5)	<b>50.3 (2)</b>	39.6	17.1	29.9
Cloud	29	58.6 (2)	60.0 (6)	38.6	16.2	<b>71.0</b>	34.9 (2)	36.6 (2)	29.9	8.8	<b>52.8</b>
Desert	54	69.3 (9)	<b>72.2 (2)</b>	48.1	23.9	43.0	38.2 (2)	<b>38.7 (2)</b>	27.2	18.8	23.4
Dense urban	73	<b>100.0 (3)</b>	<b>100.0 (3)</b>	84.1	80.4	84.2	96.4 (6)	<b>97.5 (3)</b>	62.3	65.2	65.1
Forest	69	92.2 (5)	<b>93.5 (6)</b>	53.8	33.0	63.9	67.2 (3)	<b>68.6 (6)</b>	37.7	21.4	41.5
Mountain	53	75.3 (7)	<b>75.8 (4)</b>	56.2	37.0	63.4	<b>45.5 (9)</b>	43.7 (4)	39.0	26.2	44.2
Ocean	68	87.9 (3)	<b>91.5 (4)</b>	33.7	45.0	58.8	61.1 (5)	<b>68.9 (6)</b>	17.3	23.7	28.7
Snow	55	87.6 (7)	<b>90.0 (5)</b>	46.9	35.5	71.3	42.3 (2)	<b>49.7 (2)</b>	29.1	17.1	39.9
Wetland	35	19.7 (5)	22.0 (2)	<b>31.1</b>	7.1	28.6	15.7 (9)	15.7 (2)	<b>20.9</b>	9.4	18.8
<i>ICONES-HSI</i> (all classes)	486	77.3 (5)	<b>80.3 (2)</b>	52.3	37.7	60.5	53.2 (5)	<b>56.2 (5)</b>	35.0	26.0	39.2

**Dataset:** Composed of 486 radiance images, *ICONES-HSI* is a dataset of remote sensing images collected using AVIRIS sensor. The images are organized into nine classes i.e. agriculture (50), cloud (29), desert (54), dense urban (73), forest (69), mountain (53), ocean (68), snow (55), and wetland (35). Each image measures  $300 \times 300$  pixels with  $L = 224$  spectral bands ranging from 365 nm to 2497 nm, covering both visible and near infrared regions. To eliminate the manual border corrections<sup>2</sup> in some images, we crop the image from the center so that the final images measure  $250 \times 250$  pixels.

A careful examination reveals a spectral registration error or redundancy at several wavelengths. For example, the 32th band reads 667.561 nm, but the 33th and 34th bands read 655.7923 nm and 665.5994 nm respectively. This is due to the overlapping of spectral acquisitions between the four sensors employed for AVIRIS data acquisition. As such, we remove the 32th, 33th, 96th, 97th, 160th, and 161th bands. On the other hand, contrary to common practices we keep the noisy e.g. water absorption bands to stay as faithful as possible to the measurements. This also allows us to test the resistance of our feature to noise which is a practical subject in real world applications. Hence, we consider only  $L = 218$  bands.

**Computations:** The similarity between images are ranked by Kullback-Leibler divergence between their RSDOM modeled using GMM. The same parameters as in Section VIII-A are used for the RSDOM calculation, hence a monoscale and rotation-invariant assessment with two spectral references. However, we disregard the spectral intensity information,  $\Delta W$  as it offers no help for scene recognition in this CBIR context. This is because the same scene could exhibit different brightness depending on the time of acquisition e.g. brighter if taken in noon, or darker if acquired in early morning or late evening. This induces a four-dimensional RSDOM considering just the joint distribution of  $\Delta G_1$ ,  $\Delta G_2$ ,  $\Delta G'_\Theta$ , and  $\Delta W'_\Theta$ . This demonstrates the easy adaptation of RSDOM in a given context thanks to its metrological constructions as opposed to some state-of-the-art which fuzz the spectral and spatial char-

acterization. As for GLCM, Gabor filter and LBP approaches, we consider their cross-channel versions applied on the top three principal components (PCs) in line with the common practice in the community [63]–[65]. For *ICONES-HSI*, this corresponds to an average (over all images) explained variance of 98.0%. Other than that, the same implementation and similarity measurement as in Section VIII-A are employed.

**Results:** The performances of RSDOM as well as GLCM, Gabor filter, and LBP are listed in Table IV in terms of the precision at top 10 retrieved images (P@10) and the mean average precision (MAP) over the retrieved data. It is observed that RSDOM (P@10 = 80.3%) outperforms GLCM (52.3%), Gabor filter (37.7%), and LBP (60.5%) by a considerably wide margin. Interestingly, spectral features alone (using  $\mathcal{J}_1$ ) also significantly outmatches (77.3%) the state-of-the-art considered, further confirming the efficiency of our spectral characterization. To avoid suspicion on the inefficiency of GLCM, Gabor, and LBP due to the low number of PC used, we repeat the experiment using the top 18 PCs for an average explained variance of 99.7%. The results indicate improvements only for LBP with (P@10, MAP) = (64.7%, 44.7%) which is still far below RSDOM (80.3%, 56.2%), whereas GLCM (42.3%, 28.1%) and Gabor (19.0%, 24.0%) actually perform worse.

Fig. 14 demonstrates few instances of the CBIR using RSDOM. Clearly, the performances of RSDOM is satisfying. In particular, the dense urban images (P@10 = 100%) are perfectly identified by RSDOM due to its distinct spectral (man-made objects) and spatial distribution (grid-like structure). The ocean images (91.5%) are also efficiently retrieved thanks to their highly stationary texture which is easily captured by RSDOM. The forest images (93.5%), too, are easily identified due to their signature chlorophyll reflectance which is in turn, efficiently characterized by RSDOM. Although the agriculture images (79.0%) also exhibit signature chlorophyll reflectance, they are readily differentiated from the forest images thanks to their block-like structure. This again illustrates the interest of a joint spectral-spatial texture characterization in RSDOM. However, the cloud (60.0%) and wetland images (22.0%) are poorly identified by RSDOM. This is due to the fact that cloud

<sup>2</sup>In some images, a significant number of black pixels can be seen filling the gap from the border of the acquisition as the plane track is not straight.

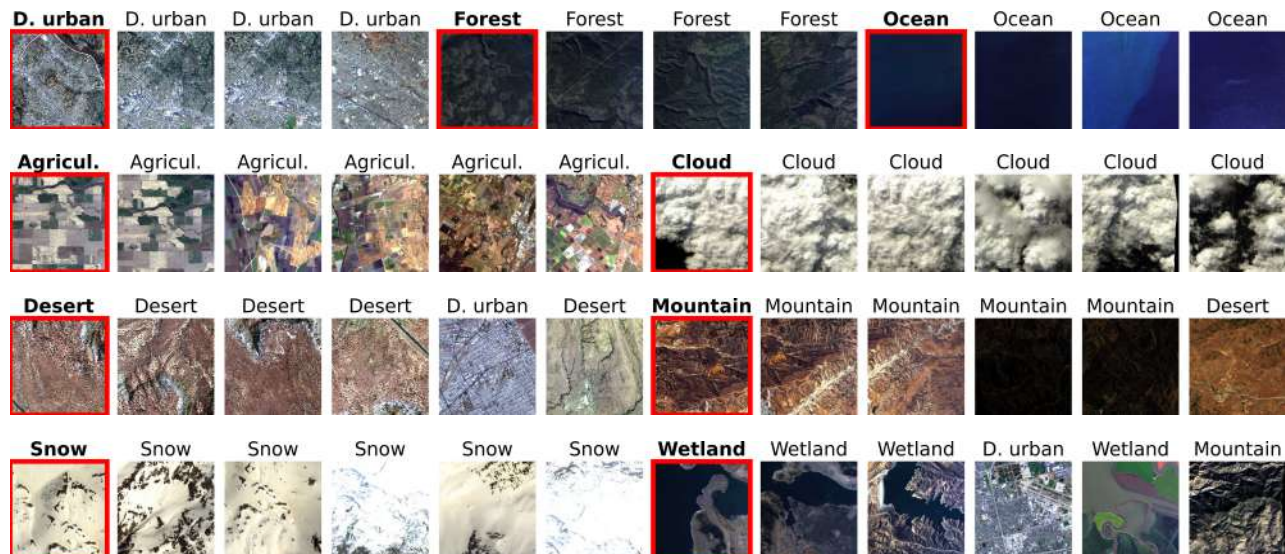


Fig. 14: Instances of image retrieval results using RSDOM on *ICONES-HSI*. Bordered in red are the query images, followed by the top retrieved images. To avoid favorability in this displayed example, the first image in each class is taken as the query.

is a highly complex texture with a large fractal dimension. Therefore, a multiscale assessment is necessary which could be done by processing RSDOM for several scales. Furthermore, in some cases land and buildings can be seen underneath the clouds, thus inducing confusion with other classes like dense urban. On the other hand, the wetland images are highly heterogeneous as they contain a large variety of textures e.g. lake, plantations, buildings, and bare soils. Such complexity could be nonetheless tackled by adapting RSDOM using a bag-of-words (BoW) model, thus just a question of optimization.

### C. Land cover classification on Salinas

In this third and final task, we assess the performance of RSDOM in a land cover classification scheme on one of the standard benchmark datasets for hyperspectral classification. Different from the previous two tasks, the efficiency of RSDOM as a local descriptor is highlighted here as the characterization is performed on pixel-wise levels.

**Dataset:** The *Salinas* dataset is a hyperspectral image measuring  $512 \times 217$  pixels with  $L = 224$  spectral bands collected using AVIRIS sensor over the *Salinas* valley, California. The spatial resolution is given by 3.7 m/pixel with various textures e.g. plantations and soils. About 49% of the pixels have been manually labeled with one of the 16 classes identified. For direct comparison with the state-of-the-art [66], [67], we follow the convention of removing the water absorption bands (224th, 154th–167th, and 108th–112th). We also remove the misregistered wavelengths, thus leaving  $L = 200$  bands.

**Computations:** The calculation of RSDOM is carried out as in Section VIII-A except that the assessment is done on  $K = 8$  directions,  $\Theta = \{0, \pi/4, \pi/2, 3\pi/4, \pi, 5\pi/4, 3\pi/2, 7\pi/4\}$  due to the small spatial support in the pixel-wise classification. However, RSDOM is not defined for a single pixel as it is formulated as a probability density function. As such, a spatial neighborhood of modest size<sup>3</sup>,  $7 \times 7$  is considered for

the five-dimensional PDF construction. In light of the very reduced number of pixels, a normal approximation ( $M = 1$ , thus a feature size of a mere 20 scalars) is employed for the statistical modeling. We compare its performance with cross-channel LBP which been shown to be superior to both GLCM and Gabor filter in Section VIII-A and VIII-B. The same implementation as before is employed for LBP by considering the top three principal components. In addition, we also incorporate results from [66], [67] for comparisons with machine learning approaches as briefly explained below:

- SVM: A support vector machine implemented with a radial basis function (RBF) kernel [66], [69].
- CCF: A canonical correlation forest with 200 trees belonging to the decision tree ensemble methods [66], [70].
- CNN: A VGG-like network with spectral attention module which emphasizes on informative bands [66].
- GCN: A non-local graph convolutional network taking advantage of both labeled and unlabeled data [67].

To reduce bias, we repeat the classification 100 times. Following the same protocol as in [66], [67], 50 pixels from each class are randomly selected for training as exemplified in Fig. 15 (k). As the pixel-wise classification may be less robust to noise (thus producing the “salt and pepper” effect), some post-processing is applied to improve the spatial correlation. In this regard, we perform a minimalistic  $3 \times 3$  majority filtering. For the evaluation, we consider the mean and standard deviation of the overall accuracy (OA) i.e. the percentage of test samples correctly classified, per-class accuracy (PA) i.e. the percentage of test samples correctly classified for each class (useful for imbalanced classification as in the case of *Salinas*), average accuracy (AA) i.e. the mean of all PAs, and Kappa coefficient (KP) i.e. a more robust measure of inter-rater reliability.

**Results:** From Table V and Fig. 15 (b) - (d), it can be seen that RSDOM performs significantly better than both LBP and machine learning methods. In fact in spectral classification alone, the “Spectral” (OA=93.7%) part of RSDOM i.e.  $\mathcal{I}_1$  already overpasses both SVM (88.8%) and CCF (89.7%) as

<sup>3</sup>In statistics, the minimum sample size for normality is 30 (heuristic) [68]. The smallest (odd-sized) neighborhood respecting this constraint is  $7 \times 7$ .

TABLE V: COMPARISONS OF PER-CLASS ACCURACY, OVERALL ACCURACY (OA), AVERAGE ACCURACY (AA), AND KAPPA COEFFICIENT (KP) FOR LAND COVER CLASSIFICATION ON *Salinas*. BOLD NUMBERS INDICATE THE BEST PERFORMANCE IN EACH CLASSIFICATION (SPECTRAL/SPATIAL, POST-PROCESSED). RESULTS AVERAGED OVER 100 CLASSIFICATIONS.

Class	<i>No. pixel</i>		<i>Spectral classification</i>			<i>Spectral-spatial classification</i>				<i>Post-processed</i>		
	Train	Test	Spectral	SVM	CCF	RSDOM	LBP	CNN	GCN	Spectral+	RSDOM+	LBP+
Brocoli...1	50	1959	99.1±1.4	99.0	<b>99.5</b>	99.4±1.2	99.6±0.5	94.8	<b>99.7</b>	99.9±0.6	99.9±0.5	<b>100±0.0</b>
Brocoli...2	50	3676	98.4±1.0	99.7	<b>100</b>	98.9±0.9	97.0±1.2	<b>100</b>	99.2	99.4±0.7	<b>99.5±0.6</b>	98.4±1.2
Fallow	50	1926	98.3±1.0	98.7	<b>99.4</b>	98.2±1.0	93.8±2.0	99.6	<b>99.8</b>	<b>99.5±0.7</b>	99.3±0.7	96.9±2.0
Fallow....plow	50	1344	98.7±1.2	97.8	<b>99.3</b>	<b>99.2±0.8</b>	98.2±1.1	98.9	98.3	99.7±0.7	<b>100±0.2</b>	99.8±0.5
Fallow_smooth	50	2628	95.1±2.1	98.3	<b>98.8</b>	95.7±2.0	89.4±2.6	<b>99.8</b>	99.3	96.8±1.9	<b>97.2±1.8</b>	93.0±2.8
Stubble	50	3909	99.7±0.4	99.7	<b>99.8</b>	99.7±0.4	<b>100±0.1</b>	99.7	99.8	100±0.1	<b>100±0.0</b>	<b>100±0.0</b>
Celery	50	3529	98.8±1.2	99.5	<b>99.7</b>	99.0±1.2	95.9±1.2	<b>99.7</b>	99.0	<b>99.7±0.6</b>	99.7±0.7	98.0±1.1
Grapes...	50	11221	<b>83.7±2.8</b>	70.4	67.6	84.2±3.0	67.0±3.1	<b>84.3</b>	79.1	87.7±3.0	<b>88.0±3.2</b>	75.9±3.8
Soil_vinyard...	50	6153	97.5±1.3	98.6	<b>99.2</b>	97.9±1.1	80.2±3.1	<b>98.4</b>	97.7	98.9±0.9	<b>99.0±0.9</b>	90.5±3.4
Corn...weeds	50	3228	<b>96.7±1.4</b>	93.7	93.8	<b>97.4±1.3</b>	89.8±2.1	95.1	95.0	98.6±1.1	<b>98.9±1.0</b>	96.1±2.0
Lettuce...4wk	50	1018	<b>99.0±1.6</b>	94.7	95.9	<b>99.2±1.3</b>	97.3±1.5	98.8	94.6	99.5±0.9	<b>99.6±0.9</b>	99.3±1.1
Lettuce...5wk	50	1877	99.5±0.4	99.9	<b>100</b>	99.7±0.3	94.8±2.0	99.6	<b>100</b>	<b>100±0.0</b>	<b>100±0.0</b>	97.8±1.8
Lettuce...6wk	50	866	<b>99.5±0.4</b>	97.8	98.2	99.5±0.4	98.7±1.1	<b>100</b>	99.0	99.8±0.3	<b>99.9±0.2</b>	99.6±0.9
Lettuce...7wk	50	1020	<b>99.5±0.8</b>	97.4	96.9	99.6±0.7	98.1±1.2	<b>99.9</b>	99.4	99.8±0.7	<b>99.8±0.6</b>	99.5±0.9
Vin._untrained	50	7218	<b>87.4±2.9</b>	71.5	80.8	<b>87.8±2.8</b>	82.2±3.3	79.4	84.3	91.6±2.8	<b>92.1±2.7</b>	90.5±3.3
Vin....trellis	50	1757	98.1±1.0	<b>98.2</b>	<b>98.2</b>	<b>98.7±0.9</b>	94.5±2.0	96.9	98.0	99.0±0.9	<b>99.4±0.7</b>	98.4±1.4
	800	53329										
OA			<b>93.7±0.6</b>	88.8	89.7	<b>94.0±0.6</b>	86.0±0.8	93.0	92.5	95.8±0.6	<b>96.0±0.7</b>	91.4±0.9
AA	-	-	<b>96.8±0.3</b>	94.7	95.4	<b>97.1±0.3</b>	92.3±0.5	96.6	96.4	98.1±0.3	<b>98.3±0.3</b>	95.8±0.5
KP ×100			<b>93.0±0.6</b>	87.6	88.6	<b>93.4±0.7</b>	84.5±0.8	92.2	91.6	95.3±0.7	<b>95.5±0.7</b>	90.5±0.9

well as the spectral-spatial methods i.e. LBP (86.0%), CNN (93.0%), and GCN (92.5%). This shows the high efficiency of our spectral characterization which is based on spectral differences with two references. Such achievement is further surpassed through the incorporation of spatial information by RSDOM which records an OA of 94.0%. Together with results in Section VIII-A and VIII-B, the importance of metrological considerations in feature extraction is clear. When metrology is strictly observed at all levels, even low-level features can be as effective as complex machine learning approaches.

Referring to Fig. 15 (i) and (j), we discuss the four main causes of misclassification by RSDOM. The first reason is the high texture nonstationarity in “grapes\_untrained” as exemplified in region (I) which is clearly of completely different texture. In fact, “vinyard\_untrained” is often confused with “grapes\_untrained” although RSDOM produces the best discrimination compared to other methods. The second reason is the existence of foreign objects in the labeled data as exemplified by the black dot in “fallow\_smooth” in region (II). As a result, the pixels around are greatly misclassified. The third reason is pixel mislabeling which occurs at quite many places as exemplified by the buildings labeled as “corn\_senesced\_weeds” at region (III). Evidently, the manual labeling is prone to human errors and care must be taken to avoid features “learning” these errors. The fourth reason is due to the spatial neighborhood of RSDOM which causes problems at fine boundaries as in region (IV) as a result of oversmoothing or border spillover. As solution, the neighborhood size may be reduced but this comes at a cost of less accurate statistical modeling. Nevertheless, we show that this problem can also be easily overcome by some post-processing e.g. majority filtering as shown in Fig. 15 (f) - (h).

As indicated in Table V, this improves the OA of Spectral, RSDOM, and LBP by 2.1%, 2.0%, and 5.4% respectively.

On a side note, for comparison with SVM, CCF, CNN, and GCN (of which the results are taken directly from [66], [67]) we employ the same performance evaluation protocol which also coincides with public convention [71]–[75] i.e. by excluding the training pixels as exemplified in Fig. 15 (k). However, for a *real* evaluation the entire spatial neighborhood of the training pixels as in Fig. 15 (l) should be excluded. This is due to the possible leakage of information when the neighborhood of the training and testing pixels overlap [76].

## IX. CONCLUSION

We have devised a spectral difference based texture feature using full-band processing for hyperspectral image. The formulation is based on the proposed definition of texture as joint probability of spectral and spatial distribution. The resulted feature, RSDOM is expressed as a multidimensional probability density distribution. Thanks to the metrological spectral difference calculation, the formulation is generic and applicable for images of any spectral range and number of bands. The metrological properties are further preserved via the construction of texture similarity measure using Kullback-Leibler divergence based on the probabilistic nature of RSDOM. Being highly interpretable, RSDOM is fully metrological with a direct relationship between the feature and texture.

The discriminability of RSDOM is confirmed thanks to the performances in three versatile tasks: texture classification on *HyTexiLa*, content-based image retrieval on *ICONES-HSI*, and land cover classification on *Salinas*. For the considered tasks, RSDOM overpasses GLCM, Gabor filter, LBP, SVM, and CCF and archives the same level as CNN and GCN.

## REFERENCES

- [1] P. De Bièvre, “The 2012 international vocabulary of metrology: “vim”,” *Accreditation and Quality Assurance*, vol. 17, 2012.
- [2] R. J. Chu, N. Richard, C. F-Maloigne, and J. Y. Hardeberg, “A metrological measurement of texture in hyperspectral images using relocated spectral difference occurrence matrix,” in *2019 IEEE International Conference on Image Processing (ICIP)*, 2019, pp. 3133–3137.

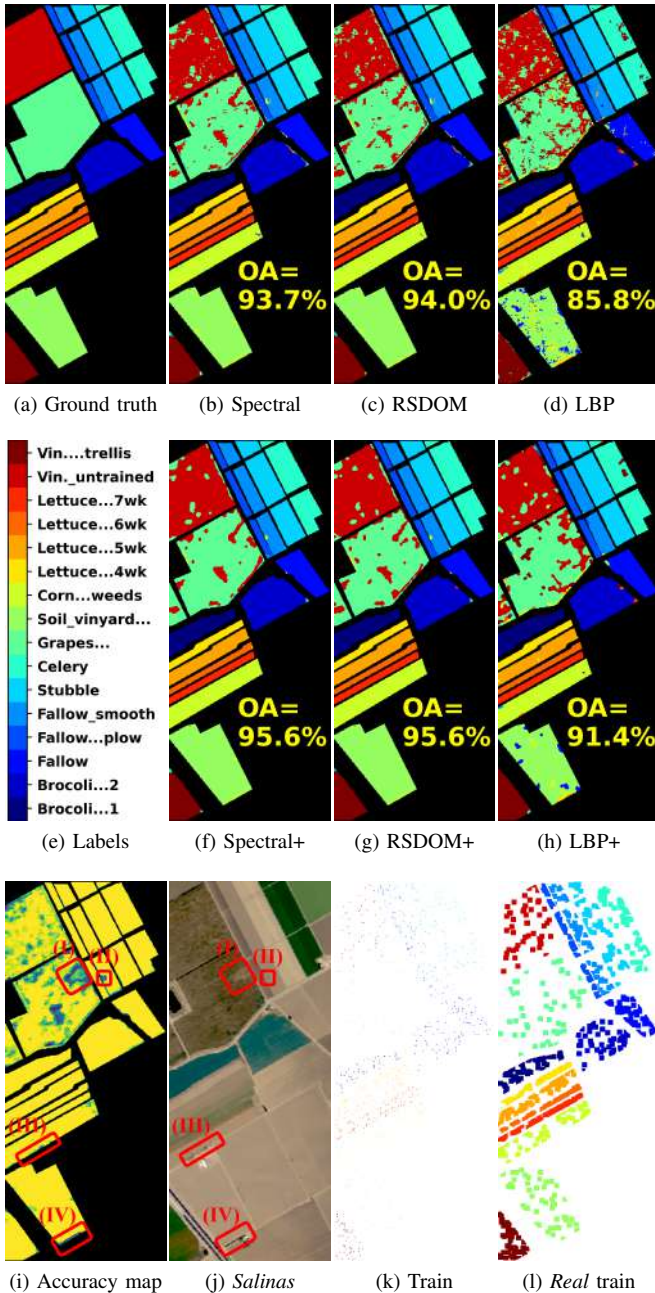


Fig. 15: (a) - (h): A comparison of performances on land cover classification on *Salinas*. In black are the unlabeled pixel which are not classified. The “+” sign indicates post-processed. (i) - (j): Accuracy map (averaged over 100 classifications) for RSDOM, with classified and misclassified pixels in yellows and blues respectively. Marked in red are the misclassification instances referred to in the discussion. (k) - (l) The train pixels for (a) - (h) and their  $7 \times 7$  neighborhood (*real train pixels*).

- [3] R. J. Chu, N. Richard, F. Ghorbel, C. Fernandez-Maloigne, and J. Y. Hardeberg, “A metrological framework for hyperspectral texture analysis using relative spectral difference occurrence matrix,” in *2019 10th Workshop on Hyperspectral Imaging and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, 2019, pp. 1–5.
- [4] A. Ledoux, N. Richard, A. S. Capelle-Laizé, H. Deborah, and C. Fernandez-Maloigne, “Toward a full-band texture features for spectral images,” in *2014 IEEE International Conference on Image Processing (ICIP)*, 2014, pp. 708–712.
- [5] A. Mirhashemi, “Introducing spectral moment features in analyzing the spectex hyperspectral texture database,” *Machine Vision and Applications*, vol. 29, pp. 415–432, 2018.
- [6] M. Uzair, A. Mahmood, and A. Mian, “Hyperspectral face recognition with spatio-spectral information fusion and pls regression,” *IEEE Trans. Image Process.*, vol. 24, no. 3, pp. 1127–1137, 2015.
- [7] G. Taşkın, H. Kaya, and L. Bruzzone, “Feature selection based on high dimensional model representation for hyperspectral images,” *IEEE Trans. Image Process.*, vol. 26, no. 6, pp. 2918–2928, 2017.
- [8] Y. Cai, X. Liu, and Z. Cai, “Bs-nets: An end-to-end framework for band selection of hyperspectral image,” *IEEE Trans. Geosci. and Remote Sens.*, vol. 58, no. 3, pp. 1969–1984, 2020.
- [9] G. Mercier and M. Lennon, “On the characterization of hyperspectral texture,” in *IEEE International Geoscience and Remote Sensing Symposium*, 2002, vol. 5, pp. 2584–2586.
- [10] M. Shi and G. Healey, “Hyperspectral texture recognition using a multiscale opponent representation,” *IEEE Trans. Geosci. and Remote Sens.*, vol. 41, no. 5, pp. 1090–1095, 2003.
- [11] L. Huo and P. Tang, “Spectral and spatial classification of hyperspectral data using svms and gabor textures,” in *2011 IEEE International Geoscience and Remote Sensing Symposium*, 2011, pp. 1708–1711.
- [12] O. Rajadell, P. García-Sevilla, and F. Pla, “Textural features for hyperspectral pixel classification,” in *Pattern Recognit. and Image Analysis*, Berlin, Heidelberg, 2009, pp. 208–216, Springer Berlin Heidelberg.
- [13] W. Li, C. Chen, H. Su, and Q. Du, “Local binary patterns and extreme learning machine for hyperspectral imagery classification,” *IEEE Trans. Geosci. and Remote Sens.*, vol. 53, no. 7, pp. 3681–3693, 2015.
- [14] H. A. Khan, S. Mihoubi, B. Mathon, J. B. Thomas, and J. Y. Hardeberg, “Hytexila: High resolution visible and near infrared hyperspectral texture images,” *Sensors*, vol. 18, no. 7, 2018.
- [15] W. Sun and Q. Du, “Hyperspectral band selection: A review,” *IEEE Geosci. Remote Sens. Magazine*, vol. 7, no. 2, pp. 118–139, 2019.
- [16] E. Esser, M. Moller, S. Osher, G. Sapiro, and J. Xin, “A convex model for nonnegative matrix factorization and dimensionality reduction on physical space,” *IEEE Trans. Image Process.*, vol. 21, no. 7, pp. 3239–3252, 2012.
- [17] K. N. Quinn, C. B. Clement, F. D. Bernardis, M. D. Niemack, and J. P. Sethna, “Visualizing probabilistic models and data with intensive principal component analysis,” *Proceedings of the National Academy of Sciences*, vol. 116, no. 28, pp. 13762–13767, 2019.
- [18] F. Tsai and J. Lai, “Feature extraction of hyperspectral image cubes using three-dimensional gray-level cooccurrence,” *IEEE Trans. Geosci. and Remote Sens.*, vol. 51, no. 6, pp. 3504–3513, 2013.
- [19] S. Jia, J. Hu, J. Zhu, X. Jia, and Q. Li, “Three-dimensional local binary patterns for hyperspectral imagery classification,” *IEEE Trans. Geosci. and Remote Sens.*, vol. 55, no. 4, pp. 2399–2413, 2017.
- [20] T. C. Bau, S. Sarkar, and G. Healey, “Hyperspectral region classification using a three-dimensional gabor filterbank,” *IEEE Trans. Geosci. and Remote Sens.*, vol. 48, no. 9, pp. 3457–3464, 2010.
- [21] X. Guo, X. Huang, and L. Zhang, “Three-dimensional wavelet texture feature extraction and classification for multi/hyperspectral imagery,” *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 12, pp. 2183–2187, 2014.
- [22] R. M. Haralick, K. Shanmugam, and I. Dinstein, “Textural features for image classification,” *IEEE Trans. Syst., Man, Cybern.*, vol. 3, no. 6, pp. 610–621, 1973.
- [23] T. Ojala, M. Pietikäinen, and T. Mäenpää, “Multiresolution gray-scale and rotation invariant texture classification with local binary patterns,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 7, pp. 971–987, 2002.
- [24] J. M. Coggins and A. K. Jain, “A spatial filtering approach to texture analysis,” *Pattern Recognit. Letters*, vol. 3, no. 3, pp. 195 – 203, 1985.
- [25] S. G. Mallat, “A theory for multiresolution signal decomposition: the wavelet representation,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, no. 7, pp. 674–693, 1989.
- [26] M. Haindl and S. Mikes, “Unsupervised texture segmentation using multispectral modelling approach,” in *18th International Conference on Pattern Recognition (ICPR’06)*, 2006, vol. 2, pp. 203–206.



- [27] G. G. Hazel, "Multivariate gaussian mrf for multispectral scene segmentation and anomaly detection," *IEEE Trans. Geosci. and Remote Sens.*, vol. 38, no. 3, pp. 1199–1211, 2000.
- [28] C. Panigrahy, A. Garcia-Pedrero, A. Seal, D. Rodríguez-Esparragón, N. H. Mahato, and C. Gonzalo-Martín, "An approximated box height for differential-box-counting method to estimate fractal dimensions of gray-scale images," *Entropy*, vol. 19, no. 10, 2017.
- [29] C. Panigrahy, A. Seal, N. H. Mahato, and D. Bhattacharjee, "Differential box counting methods for estimating fractal dimension of gray-scale images: A survey," *Chaos, Solitons & Fractals*, vol. 126, pp. 178–202, 2019.
- [30] M. Haindl and S. Mikes, "Unsupervised texture segmentation using multispectral modelling approach," in *18th International Conference on Pattern Recognition (ICPR'06)*, 2006, vol. 2, pp. 203–206.
- [31] R. Richtl and M. Haindl, "Dynamic texture similarity criterion," in *2018 24th International Conference on Pattern Recognition (ICPR)*, 2018, pp. 904–909.
- [32] D. Bhattacharjee and H. Roy, "Pattern of local gravitational force (plgf): A novel local image descriptor," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 43, no. 2, pp. 595–607, 2021.
- [33] C. Palm, "Color texture classification by integrative co-occurrence matrices," *Pattern Recognit.*, vol. 37, no. 5, pp. 965–976, 2004.
- [34] M. N. Do and M. Vetterli, "Wavelet-based texture retrieval using generalized gaussian density and kullback-leibler distance," *IEEE Trans. Image Process.*, vol. 11, no. 2, pp. 146–158, 2002.
- [35] S. K. Choy and C. S. Tong, "Statistical wavelet subband characterization based on generalized gamma density and its application in texture retrieval," *IEEE Trans. Image Process.*, vol. 19, no. 2, pp. 281–289, 2010.
- [36] S. Cui and M. Datcu, "Statistical wavelet subband modeling for multi-temporal sar change detection," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 4, pp. 1095–1109, 2012.
- [37] B. Julesz, "Visual pattern discrimination," *IRE Transactions on Information Theory*, vol. 8, no. 2, pp. 84–92, 1962.
- [38] B. Julesz, E. Gilbert, and J. Victor, "Visual discrimination of textures with identical third-order statistics," *Biological cybernetics*, vol. 31, pp. 137–40, 1978.
- [39] F. Tomita, Y. Shirai, and S. Tsuji, "Description of textures by a structural analysis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 4, no. 2, pp. 183–191, 1982.
- [40] G. Van de Wouwer, P. Scheunders, and D. Van Dyck, "Statistical texture characterization from discrete wavelet representations," *IEEE Trans. Image Process.*, vol. 8, no. 4, pp. 592–598, 1999.
- [41] P. Diaconis and D. Freedman, "On the statistics of vision: The Julesz conjecture," *Journal of Mathematical Psychology*, vol. 24, no. 2, pp. 112–138, 1981.
- [42] J. I. Yellott, "Implications of triple correlation uniqueness for texture statistics and the Julesz conjecture," *J. Opt. Soc. Am. A*, vol. 10, no. 5, pp. 777–793, 1993.
- [43] S. Kullback and R. A. Leibler, "On info. and suff.," *Annals of Mathematical Statistics*, vol. 22, no. 1, pp. 79–86, 1951.
- [44] H. Deborah, *Towards Spectral Mathematical Morphology*, Ph.D. thesis, Norwegian Univ. of Science and Technology and Univ. of Poitiers, 2016.
- [45] M. Unser, "Sum and difference histograms for texture classification," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 8, no. 1, pp. 118–125, 1986.
- [46] T. Ojala, K. Valkealahti, E. Oja, and M. Pietikäinen, "Texture discrimination with multidimensional distributions of signed gray-level differences," *Pattern Recognit.*, vol. 34, no. 3, pp. 727–739, 2001.
- [47] J. S. Weszka, C. R. Dyer, and A. Rosenfeld, "A comparative study of texture measures for terrain classification," *IEEE Trans. Syst., Man, Cybern.*, vol. 6, no. 4, pp. 269–285, 1976.
- [48] H. Deborah, N. Richard, and J. Y. Hardeberg, "A comprehensive evaluation of spectral distance functions and metrics for hyperspectral image processing," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 6, pp. 3224–3234, 2015.
- [49] N. Richard, D. Helbert, C. Olivier, and M. Tamisier, "Pseudo-divergence and bidimensional histogram of spectral differences for hyperspectral image processing," *Journal of Imaging Science and Technology*, vol. 60, no. 5, pp. 1–13, 2016.
- [50] R. H. Yuhas, A. F. H. Goetz, and J. W. Boardman, "Discrimination among semi-arid landscape endmembers using the spectral angle mapper (sam) algorithm," in *Proc. Summaries 3rd Annu. JPL Airborne Geosci. Workshop*, 1992, pp. 147–149.
- [51] C. I. Chang, "Spectral information divergence for hyperspectral image analysis," in *IEEE 1999 International Geoscience and Remote Sensing Symposium. IGARSS'99*, 1999, vol. 1, pp. 509–511.
- [52] R. V. Foutz and R. C. Srivastava, "The performance of the likelihood ratio test when the model is incorrect," *The Annals of Statistics*, vol. 5, no. 6, pp. 1183–1194, 1977.
- [53] S. Eguchi and J. Copas, "Interpreting kullback-leibler divergence with the neyman-pearson lemma," *Journal of Multivariate Analysis*, vol. 97, no. 9, pp. 2034 – 2040, 2006.
- [54] M. Basseville, "Divergence measures for statistical data processing—an annotated bibliography," *Signal Processing*, vol. 93, no. 4, pp. 621–633, 2013.
- [55] S. Simić, S. S. Alzaid, and H. Aydi, "On the symmetrized s-divergence," *Open Mathematics*, vol. 18, no. 1, pp. 378–385, 2020.
- [56] J. R. Hershey and P. A. Olsen, "Approximating the kullback leibler divergence between gaussian mixture models," in *2007 IEEE International Conference on Acoustics, Speech and Signal Processing - ICASSP '07*, 2007, vol. 4, pp. IV–317–IV–320.
- [57] N. Richard, M. Ivanovici, and A. Bony, "Toward a metrology for non-uniform surface using the complexity notion," in *4th CIE Expert Symposium on Colour and Visual Appearance*, 2016, pp. 40–50.
- [58] T. Mäenpää and M. Pietikäinen, *Texture analysis with local binary patterns*, pp. 197–216, World Scientific Publishing Co. Pte. Ltd., 2005.
- [59] M. Petrou and P. G. Sevilla, *Image Processing: Dealing with Texture*, chapter 3.8, p. 278, John Wiley & Sons, Ltd, Sussex, England, 2006.
- [60] M. Hauta-Kasari, J. Parkkinen, T. Jaaskelainen, and R. Lenz, "Generalized co-occurrence matrix for multispectral texture analysis," in *Proceedings of 13th International Conference on Pattern Recognition*, 1996, vol. 2, pp. 785–789.
- [61] B. Kumar and O. Dikshit, "Spectral-spatial classification of hyperspectral imagery based on moment invariants," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 6, pp. 2457–2463, 2015.
- [62] M. J. Swain and D. H. Ballard, "Color indexing," *International Journal of Computer Vision*, vol. 7, no. 1, pp. 11–32, 1991.
- [63] H. Su, B. Zhao, Q. Du, P. Du, and Z. Xue, "Multifeature dictionary learning for collaborative representation classification of hyperspectral imagery," *IEEE Trans. Geosci. and Remote Sens.*, vol. 56, no. 4, pp. 2467–2484, 2018.
- [64] S. Jia, K. Wu, J. Zhu, and X. Jia, "Spectral-spatial gabor surface feature fusion approach for hyperspectral imagery classification," *IEEE Trans. Geosci. and Remote Sens.*, vol. 57, no. 2, pp. 1142–1154, 2019.
- [65] W. Huang, Y. Huang, H. Wang, Y. Liu, and H. J. Shim, "Local binary patterns and superpixel-based multiple kernels for hyperspectral image classification," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 4550–4563, 2020.
- [66] L. Mou and X. X. Zhu, "Learning to pay attention on spectral domain: A spectral attention module-based convolutional network for hyperspectral image classification," *IEEE Trans. Geosci. and Remote Sens.*, vol. 58, no. 1, pp. 110–122, 2020.
- [67] L. Mou, X. Lu, X. Li, and X. X. Zhu, "Nonlocal graph convolutional networks for hyperspectral image classification," *IEEE Trans. Geosci. and Remote Sens.*, vol. 58, no. 12, pp. 8246–8257, 2020.
- [68] S. Allende-Alonso, C. N. Bouza-Herrera, S. E. H. Rizvi, and J. M. Sautto-Vallejo, "Big data and the central limit theorem: A statistical legend," *Investigación operacional*, vol. 40, no. 1, pp. 112–, 2019.
- [69] B. E. Boser, I. M. Guyon, and V. N. Vapnik, "A training algorithm for optimal margin classifiers," in *Proceedings of the Fifth Annual Workshop on Computational Learning Theory*, 1992, p. 144–152.
- [70] T. Rainforth and F. Wood, "Canonical correlation forests," unpublished, 2015.
- [71] H. Lee and H. Kwon, "Going deeper with contextual cnn for hyperspectral image classification," *IEEE Trans. Image Process.*, vol. 26, no. 10, pp. 4843–4855, 2017.
- [72] W. Li, G. Wu, F. Zhang, and Q. Du, "Hyperspectral image classification using deep pixel-pair features," *IEEE Trans. Geosci. and Remote Sens.*, vol. 55, no. 2, pp. 844–853, 2017.
- [73] M. Zhang, W. Li, and Q. Du, "Diverse region-based cnn for hyperspectral image classification," *IEEE Trans. Image Process.*, vol. 27, no. 6, pp. 2623–2634, 2018.
- [74] L. Zhu, Y. Chen, P. Ghamisi, and J. A. Benediktsson, "Generative adversarial networks for hyperspectral image classification," *IEEE Trans. Geosci. and Remote Sens.*, vol. 56, no. 9, pp. 5046–5063, 2018.
- [75] K. Tan, F. Wu, Q. Du, P. Du, and Y. Chen, "A parallel gaussian-bernoulli restricted boltzmann machine for mining area classification with hyperspectral imagery," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 2, pp. 627–636, 2019.
- [76] J. Nalepa, M. Myller, and M. Kawulok, "Validating hyperspectral image segmentation," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 8, pp. 1264–1268, 2019.