

Supervised Classification of Very High Resolution Optical Images Using Wavelet-Based Textural Features

Olivier Regniers, Lionel Bombrun, Virginie Lafon, Christian Germain

▶ **To cite this version:**

Olivier Regniers, Lionel Bombrun, Virginie Lafon, Christian Germain. Supervised Classification of Very High Resolution Optical Images Using Wavelet-Based Textural Features. IEEE Transactions on Geoscience and Remote Sensing, 2016, 54 (6), pp.3722-3735. 10.1109/TGRS.2016.2526078 . hal-01316398

HAL Id: hal-01316398
https://hal.science/hal-01316398

Submitted on 17 May 2016

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.

Supervised classification of Very High Resolution optical images using wavelet-based textural features

Olivier Regniers, Lionel Bombrun, Virginie Lafon, and Christian Germain

***Abstract*—In this paper, we explore the potentialities of using wavelet-based multivariate models for the classification of Very High Resolution optical images. A strategy is proposed to apply these models in a supervised classification framework. This strategy includes a Content-Based Image Retrieval analysis applied on a texture database prior to the classification in order to identify which multivariate model performs the best in the context of application. Once identified, the best models are further applied in a supervised classification procedure by extracting texture features from a learning database as well as from regions obtained by a pre-segmentation of the image to classify. The classification is then operated according to the decision rules of the chosen classifier. The use of the proposed strategy is illustrated in two real case applications using Pléiades panchromatic images: the detection of vineyards and the detection of cultivated oyster fields. In both cases, at least one of the tested multivariate models displays higher classification accuracies than Gray Level Co-occurrence Matrix descriptors. Its high adaptability and the low number of parameters to be set are other advantages of the proposed approach.**

***Index Terms*—Texture, wavelet, multivariate models, very high resolution, optical image, classification.**

I. Introduction

WITH the launch of more than ten Very High Resolution (VHR) optical satellites in the past 15 years (QuickBird, GeoEYE, WorldView, Pléiades, ...), satellite image data of metric and submetric resolution became increasingly available. The level of details provided in such data enables to distinguish geometric structures only observable through their spectral properties at coarser resolutions. These geometric structures can be associated with a regular spatial organisation specific to particular types of land covers. In agricultural landscapes, this is the case of cereal crops, orchards and vineyards which typically display a periodic row structure visible in VHR image data. Because of forest management practices, young tree stands in cultivated forests may also feature specific spatial patterns. Similarly in urban areas, the juxtaposition of buildings can generate a specific framework. These specific spatial patterns can hence be exploited to detect such land covers and improve the classification of VHR optical image data.

Many studies adressed this challenge by considering patterns observed in the landscape as textures. The main objective of texture-based analysis is to explore the local spatial dependencies observed between neighboring pixels in the image. This analysis generally leads to the extraction of a small-sized set of features that can be further used in a classifier. Various approaches were proposed in the literature to represent textures for the classifcation of VHR image data. Among these approaches, the Gray Level Co-occurrence Matrix (GLCM) initially proposed in [1] is still very popular within the remote sensing community. In many publications, texture descriptors derived from GLCMs were successfully used for various remote sensing applications, e.g. the classification of urban areas [2], [3], the mapping of forest species [4], [5], the estimation of forest structure variables in mono-specific forests [6], [7], [8] and the classification of agricultural land covers [9], [10]. Rather than directly characterizing the texture in the image domain as it is the case with GLCM, other authors suggested to proceed with the texture analysis in a transformed domain of the original data by applying filter banks. For example, texture features extracted by applying scale and orientation selective Gabor filters were proposed in [11] to map hedgerows in rural landscapes. An unsupervised segmentation algorithm based on Gabor filters was also introduced in [12] for the detection of vineyards. In the same way as Gabor filters, wavelet filters also offer a multi-scale and multi-orientation framework for the texture analysis. Features such as energy and entropy [13] or GLCM descriptors [14] can be extracted from each wavelet subband to characterize the texture in this transformed domain. In another common approach, probabilistic models are used in the image domain to describe local spatial dependencies and further characterize the textural information. Markov Random Fields, known for their use in the regularization of labeled image, can be modeled with these probabilistic distributions for the classification of VHR remote sensing data [15], [16]. Finally, rather than relying on pre-defined texture features, the increasingly popular deep-learning algorithms efficiently detect patterns in images through unsupervised or semi-supervised feature learning in a deep neural network architecture with many applications in remote sensing data [17], [18].

In the past few years, texture features derived from the wavelet domain gained in popularity. The wavelet domain indeed simplifies the way textures can be apprehended as it offers a multi-scale and multi-orientation framework in which each wavelet subband contains a different part of the frequency content of the original image. Low frequencies are discarded and only high frequencies are investigated during the texture analysis. Recently, new approaches emerged to analyze textures by relying on the combined use of probabilistic

This work was supported by the Aquitaine Regional Council and CNES (Centre National d'Études Spatiales) through the TOSCA project STELLA THR. O. Regniers, L. Bombrun and C. Germain are with the Laboratoire IMS, Université de Bordeaux, Talence, e-mail: firstname.lastname@ims-bordeaux.fr. Virginie Lafon is with Géo-Transfert, e-mail: virginie.lafon@u-bordeaux.fr

models and wavelet transform. Probabilistic models used in the wavelet domain can be more simple than in the image domain as unimodal distributions are often associated with wavelet coefficients. This technique consists in modeling the distribution of wavelet coefficients with a probability density function (PDF) whose parameters are estimated according to the maximum likelihood principle [19]. These parameters are determined for each high frequency subband of the wavelet decomposition and can be further used as a textural signature. Univariate models such as the Generalized Gaussian Distribution [20] or the Generalized Gamma Distribution [21] were for example proposed to represent the marginal distribution of wavelet coefficients. However, univariate models do not fully describe the statistical behavior of the wavelet coefficients as they do not include informations relative to the spatial dependencies between these coefficients. To achieve this, multivariate models were later introduced to represent the joint distribution of the wavelet coefficients. In this case, all the wavelet coefficients contained in a defined neighborhood within the subband are considered as a local observation of the spatial dependency. All these observations are clustered in a random vector whose distribution can be further modeled using multivariate PDF. The multivariate Gaussian model, the Multivariate Generalized Gaussian Distribution (MGGD) [22], models based on Spherically Invariant Random Vectors (SIRV) [23] or copula-based models [24] are among the most frequently studied models in this context. The classification performances of these multivariate models were mostly investigated in a content-based image retrieval framework applied on natural texture databases such as VisTex [25] or Brodatz [26] and showed greater retrieval performances than univariate models. Despite the growing interest of the image processing and computer vision communities for this approach, the use of wavelet-based multivariate modeling has not yet been extensively explored for the processing of VHR optical remote sensing data. In a recent work, we demonstrated that wavelet-based multivariate models, in particular SIRV-based models, showed promising performances for the classification of forest stand age in mono-specific cultivated forests [27].

The potential of the use of multivariate models in the wavelet domain is already known in the image processing and computer vision communities [28]. The main contribution of this paper is to demonstrate that such texture analysis approaches are also suitable for the supervised classification of textured soil occupations in VHR optical remote sensing data. Moreover, we propose a complete strategy to apply such models in the context of the classification of VHR optical satellite data. This strategy consists in two steps. First, a content based image retrieval system is used to identify the best probabilistic models to be considered in the context of application. Once identified, the best models are used in a region-wise supervised classification procedure applied on a pre-partitioned image. From a more practical point of view, the main objective of this paper is also to highlight the universality of the proposed strategy which can easily be adapted to various thematic applications with a limited parameters to be set. To this aim, texture-based classification results are presented and discussed for two application examples: the detection of vineyards and the detection of cultivated oyster fields.

The paper is organized as follows. In section II, a more thorough insight on the wavelet transform and the probabilistic modeling of wavelet subbands is given. The proposed strategy for the application of wavelet-based texture features is detailed in section III. This third section is divided in two parts. In the first part, the content-based image retrieval framework is presented with a focus on similarity measurements between multivariate models. The proposed supervised classification procedure is described in the second part. In the following section, texture-based classification results are discussed and compared with a reference approach using GLCM descriptors. Conclusions and perspectives are finally presented in the last section.

II. Wavelet-Based Multivariate Modeling

Wavelet transforms provide a multi-resolution analysis framework in which the image is decomposed in a series of subband each containing an orientation and scale specific part of the frequency content of the image. This multi-resolution scheme of analysis is convenient for texture analysis as the pattern constituting the texture can be composed of elements of variable sizes. The most standard wavelet decomposition used for image filtering is the Discrete Wavelet Transform (DWT) initially proposed by Mallat [20]. This decomposition consists in applying combinations of oriented one-dimensional high-pass (H) and low-pass (L) filters to the image resulting in four subbands (three orientation specific high-pass filtered subbands (HH, HL, LH) and one low-pass filtered subband (LL)) subsampled by a factor 2 (dyadic wavelet) at each scale of the decomposition. A subsequent scale is obtained applying the same combinations of filters to the LL subband. The DWT has the advantage of producing independent orthogonal subbands without redundancy of information. Nevertheless, the DWT is limited in two points: the directional selectivity (only three main orientations can be considered) and the shift variance [29]. Other configurations of the wavelet decomposition which overcome these issues were later proposed. In wavelet packets [30], the decomposition is applied at each scale on all the subbands of the previous scales including the high-pass filtered subbands, allowing a finer analysis of the high frequencies in the image. The stationary wavelet transform [31] produces undecimated subbands preventing issues related to shift invariance. Steerable pyramids [32] enables a shift invariant and multi-orientation decomposition with more than three main orientations. Whichever the chosen wavelet decomposition, the joint distribution of wavelet coefficients in each subband can be further modeled as illustrated in Fig. 1. The aim of this paper is not to compare these various configurations of wavelet decomposition and we mainly focus on the comparison of performances of different multivariate models applied on DWT subbands.

To conduct the modeling of the spatial dependencies within a subband, the wavelet coefficients contained in the neighborhood of the current spatial position are clustered in a random vector $\mathbf{k}$. Several multivariate models were proposed to represent the distribution of the random vector $\mathbf{k}$. We

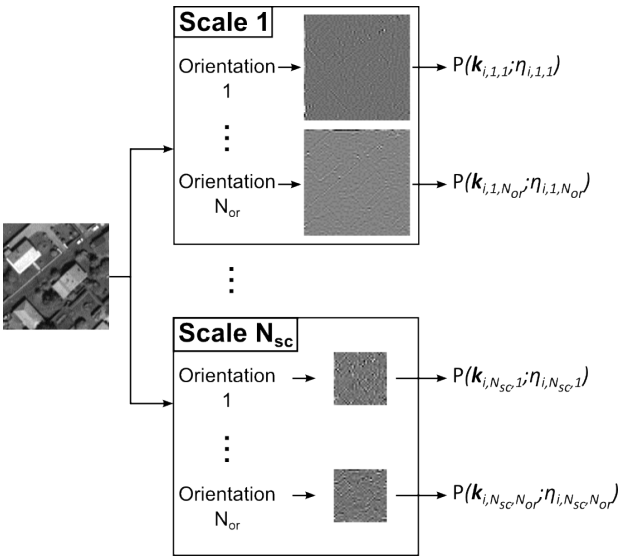


Fig. 1: Modeling of discrete wavelet subbands where $\mathbf{k}_i$ are the random vectors of observation, η are the parameters of the model P, N_{sc} and N_{or} are respectively the number of scales and orientations of the decomposition.

first focus on the well-known multivariate Gaussian model and further introduce two generalizations of this model: the SIRV-based models and the copula-based models. For each of these models, the technique used to estimate the distribution parameters is detailed in the following paragraphs. These estimated parameters η are further used as a textural signature in the classification procedure.

A. Multivariate Gaussian Model

If we assume the gaussianity of the dependency to be modeled, the observed vector $\mathbf{k}$ is a real Gaussian vector of zero mean and covariance matrix $[\mathbf{M}]$. A zero mean can be considered here as only the high frequency subbands are modeled. This stochastic Gaussian distribution is then characterized by an estimation of the covariance matrix $[\mathbf{M}]$ using the Sample Covariance Matrix estimator (SCM):

$$[\hat{\mathbf{M}}]_{SCM} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{k}_i\mathbf{k}_i' \tag{1}$$

where N is the number of observations and $\mathbf{k}_i'$ is the transposed vector of $\mathbf{k}_i$. Hence, the textural signature obtained from this model contains an estimation of the covariance matrix for each subband of the decomposition. However, it is well known that Gaussian models are not flexible enough to capture the "heavy-tail" behaviour generally observed in the distribution of wavelet coefficients [20]. Other multivariate models offering more adjustement flexibility were therefore proposed as generalizations of the Gaussian case. The SIRV-based and copula-based models are among such generalizations.

B. SIRV-based Models

In SIRV-based multivariate models, the generalization of the Gaussian case is obtained as follows. The observed vector $\mathbf{k}$ is decomposed as $\mathbf{k} = \sqrt{\tau}\mathbf{z}$ where $\mathbf{z}$ is a Gaussian vector of zero mean and covariance matrix $[\mathbf{M}]$ and τ a positive random scalar variable called the multiplier, τ and $\mathbf{z}$ being assumed independent. The multiplier τ enables to modulate the distribution of the Gaussian vector $\mathbf{z}$ to better fit the heavy-tail behaviour of the distribution of wavelet coefficients. The PDF of the random variable τ is not specified in the definition of SIRV models and various univariate models derived from the Pearson system of distributions were proposed to represent the multiplier [23]. In this paper, we will focus mainly on two particular configurations of the SIRV models. First, when considering a deterministic multiplier τ, a Gaussian-like distribution is produced (SIRVgauss). Here, the multiplier τ varies from one sample $\mathbf{k}_i$ to another and can be represented by a Dirac [33]. Second, for an Inverse Gamma distributed multiplier τ, the SIRV vector $\mathbf{k}$ follows a $\mathcal{G}^0$ distribution (SIRVg0) [34].

The SIRV representation is not unique and a normalization condition is necessary. Classicaly, the normalization is imposed on the trace of the covariance matrix, set here to p the dimension of the target vector $\mathbf{k}$. An approximated maximum likelihood estimator of the normalized covariance matrix is the solution of the following recursive equation known as the "Fixed Point" estimator (FP) [35]:

$$[\hat{\mathbf{M}}]_{FP} = f([\hat{\mathbf{M}}]_{FP}) = \frac{p}{N}\sum_{i=1}^{N}\frac{\mathbf{k}_i\mathbf{k}_i'}{\mathbf{k}_i'[\hat{\mathbf{M}}]_{FP}^{-1}\mathbf{k}_i} \quad \text{with } Tr([\hat{\mathbf{M}}]_{FP}) = p \tag{2}$$

The existence and the uniqueness of the Fixed Point estimator as well as the convergence of the recursive algorithm whatever the initialization was established in [36]. When used in SIRV configurations, the FP estimator has been proven to be more robust than the SCM estimator used for the multivariate Gaussian model [33]. For a given covariance matrix $[\mathbf{M}]$, the maximum likelihood estimator of the multiplier τ_i is then defined by:

$$\hat{\tau}_i = \frac{\mathbf{k}_i'[\mathbf{M}]^{-1}\mathbf{k}_i}{p} \tag{3}$$

In both SIRVgauss and SIRVg0 cases, the obtained textural signature contains the FP estimator of the covariance matrix, completed in the case of the SIRVg0 model with the parameters of the univariate Inverse Gamma PDF chosen to represent τ (one shape and one scale parameter).

C. Copula-based Models

Another generalization of the Gaussian case is given by the family of copula-based models. Here, the margins of the observed p-dimensional vectors $\mathbf{k} = [k_1, \ldots, k_p]$ are first modeled independently by a chosen univariate PDF. A transformation is then applied on the modeled marginals by computing a cumulative distribution function CDF $F_1, \ldots, F_p$ resulting in a set of transformed variables $(F_1(k_1), \ldots, F_p(k_p))$. Then, a copula C is used to represent the relationship between these transformed variables. The definition of the copula does not specify the distribution associated with the CDF or the marginals and various configurations of copulas are possible

[24], [37], [38]. The advantage of copulas in comparison to SIRV models is that they enable to define a joint distribution with arbitrary marginals.

The definition of the copulas is derived from the Sklar's theorem [39] which states that given a p-dimensional random vector $\mathbf{k} = [k_1, \ldots, k_p]$ with a CDF F and continuous marginals $F_1, \ldots, F_p$, there exists a unique copula C such that:

$$F(k_1, \ldots, k_p) = C(F_1(k_1), \ldots, F_p(k_p)) \tag{4}$$

Given that the CDF F and the copula C are continuous, the joint density $f(\mathbf{k})$ can be expressed as follows:

$$f(k_1, \ldots, k_p) = c(F_1(k_1), \ldots, F_p(k_p)) \prod_{r=1}^{p} f_r(k_r) \tag{5}$$

where c is the copula density and $f_r, r = 1, \ldots, p$ are the marginal PDFs. Among the configurations of copulas, we will focus on the use of a multivariate Gamma distribution with a Gaussian copula (GGC) as proposed in [37]. In this configuration, the copula density c is modeled by a Gaussian distribution and the marginals f are modeled with univariate Gamma distribution. Similarly to SIRV models, the multivariate Gaussian part of the copula model is modulated by the univariate distribution applied to the marginals in order to better capture the non-Gaussianity of the distribution of the wavelet coefficients.

In the case of Gaussian copulas, the maximum likelihood estimation of the parameters can be done in two steps. The first consists in the estimation of the parameters $\hat{\eta} = (\hat{\eta}_1, \ldots, \hat{\eta}_p)$ of each marginal $f_1, \ldots, f_p$ independently:

$$\hat{\eta}_r = \arg\max_{\eta_r} \log \prod_{r=1}^{p} f_r(k_{r,i}, \eta_r) \tag{6}$$

The estimation of the covariance matrix $[\mathbf{M}]$ is then obtained in a second step after transforming the observed vectors $\mathbf{k}_i$ into Gaussian vectors $\mathbf{k}_i^G$. The maximum likelihood estimator of the matrix $[\mathbf{M}]$ is then defined as:

$$[\hat{\mathbf{M}}] = \frac{1}{N} \sum_{i=1}^{N} \mathbf{k}_i^G \mathbf{k}_i^{G'} \tag{7}$$

Hence, the extracted textural signature obtained with Gaussian copula models is also composed of an estimation of a covariance matrix and the parameters of the univariate distribution chosen to represent the marginals (one shape and one scale parameter in the case of the Gamma distribution). Further details on the estimation of these parameters can be found in [37], [38].

The four multivariate models described above (SCM, SIRVg0, SIRVgauss, GGC) were tested in two different contexts of VHR data classification. Before presenting these results in section IV, the strategy proposed to include these wavelet-based multivariate models in a supervised classification procedure is explained in the next section.

III. Methodology

Prior to the supervised classification, a first analysis using a Content-Based Image Retrieval framework [40] is conducted on a pre-existing texture database in order to identify which are the multivariate models capable to most efficiently distinguish the considered texture classes. The concept of distance or similarity measurement between multivariate models is inherent to the CBIR framework and is defined here below. Once identified, the best models are further used in a supervised classification procedure described in the second part of this section.

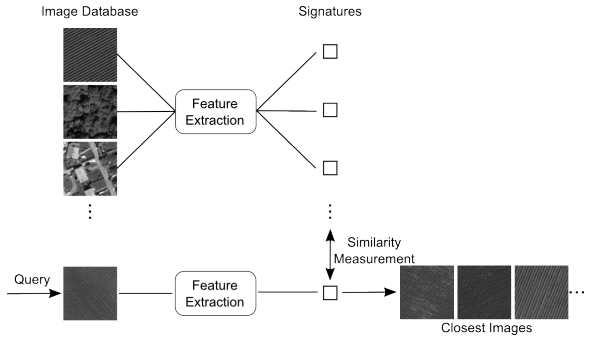


Fig. 2: Architecture of the Content-Based Image Retrieval framework (adapted from [19]).

A. Content-Based Image Retrieval

The Content-Based Image Retrieval framework (CBIR) is a universal technique aiming at the retrieval of images from a database depending on color, shape or texture features describing the visual content of these images. In this study, we focus on the use of textural features derived from wavelet-based multivariate models. This general framework is frequently used in the literature to compare the ability of probabilistic models to represent the textural content of image patches belonging to the VisTex or Brodatz databases [25], [26]. In order to apply such framework to a real practical case of image classification, a database of image patches is first created by extracting texture-wise homogeneous regions representative of the different classes of interest from the VHR optical data. Two major tasks are considered in the CBIR: the feature extraction and the similarity measurement (Fig. 2).

During the feature extraction step, the textural signature of all the patches of the database are computed according to the chosen texture analysis approach. In the case of the GLCM for example, a signature consists of a vector of descriptors derived from the co-occurrence matrix computed on one patch, whereas in the case of multivariate models, the signature contains the estimated paramaters of the modeled distribution obtained from each of the wavelet subband as described in the previous section. Once the signatures are computed, one patch is extracted from the database and considered as a query image. A similarity measurement is then calculated between the signature of the query image and all the signatures of the rest of the database, allowing to retrieve the top closest images. A textural analysis method is considered more efficient than another if amongst the top closest images, there is a higher proportion of patches belonging to the same class as the query image. This operation is repeated by considering successively each patch of the database as a query image. In the case of

models derived from wavelet subbands, it is worth noting that, by assuming the hypothesis of decorrelation between wavelet subbands, the similarity measured between two images I_1 and I_2 is equal to the sum of similarities measured between the models derived from pairs of subbands W_1 and W_2 at the same scale sc and the same orientation or of the decomposition:

$$D(I_1\|I_2) = \sum_{sc=1}^{Nsc}\sum_{or=1}^{Nor} D(W_{1sc,or}\|W_{2sc,or}) \tag{8}$$

where Nsc and Nor are respectively the number of scales and the number of orientations of the decomposition.

When using for example textural features derived from a GLCM, the notion of similarity between two images comes down to the measure of a distance between two vectors in the parameter space. A Euclidean or a Mahalanobis distance can be used in such case. For applications based on probabilistic models, Euclidean or Mahalanobis distances are not appropriate as they do not take advantage of the probabilistic nature of the parameters. The use of the Kullback-Leibler Divergence (KLD) has been widely advocated in this context [19]. The KLD between two models $P(\mathbf{k};\eta_1)$ and $P(\mathbf{k};\eta_2)$ is indeed defined according to the chosen PDF as follows:

$$KLD = D(P(\mathbf{k};\eta_1)\|P\mathbf{k};\eta_2)) = \int P(\mathbf{k};\eta_1)\log\frac{P(\mathbf{k};\eta_1)}{P(\mathbf{k};\eta_2)}d\mathbf{k} \tag{9}$$

where η refers to the model parameters. Unfortunately, a closed form of the KLD is not currently available for all multivariate models. The use of the Rao Geodesic Distance (GD) was proposed as an alternative to the KLD (e.g. [22], [41]). This distance relies on the Fisher information matrix and is defined as the length of the shortest path connecting two points in the space of the parameters of the distribution. The use of the GD is also preferred since unlike the KLD, it satisfies the triangular inequality. Besides, it was shown that the GD could lead to higher classification performances than the KLD [22], [42]. However, in the case of the GGC model, only a closed form of the KLD is currently available and is used to measure similarities for this model. The approach used to estimate similarities are described here below for the four considered multivariate models.

1) Similarity between Gaussian models: For Gaussian distributions, models are only characterized by their covariance matrix estimated in this study either with the SCM estimator for the multivariate Gaussian model (SCM) or by the FP estimator for the Gaussian SIRV model (SIRVgauss). The GD between two Gaussian models defined by the matrices $[\mathbf{M}]_1$ and $[\mathbf{M}]_2$ of dimension $p\times p$ is expressed as follows:

$$GD([\mathbf{M}]_1\|[\mathbf{M}]_2) = \sqrt{\sum_{i=1}^{p}(\ln\lambda_i)^2} \tag{10}$$

where λ_i are the eigenvalues of $[\mathbf{M}]_1^{-1}[\mathbf{M}]_2$.

2) Similarity between SIRVg0 models: Similarity measurements between two SIRVg0 models depends not only on the covariance matrix but also on the other parameters of the distribution. Various strategies were suggested to approximate the GD in this context. Here, we applied the strategy proposed by [41] in which elements of the Fisher information matrix are expressed only as a function of the two shape and scale parameters of the Inverse Gamma distribution. The GD is then approximated by assuming the geodesic coordinate functions as straight lines. Another strategy considering fixed shape parameters was also introduced in [41] for SIRVg0 models, but it was not used in this study as it was shown to be less efficient.

3) Similarity between GGC models: For copula-based models, there are currently no available closed form of the GD. A Monte-Carlo approach to approximate the KLD between two copula-based models was suggested in [24]. But, this technique often implies high computational costs. More recently, a closed form of the KLD was introduced for Gaussian copulas [38]. This expression equals to the sum of one KLD measured between two Gaussian models for the copula and p KLD between univariate models for the marginals. The symetric version of this KLD was used here for comparing two copula-based models.

In order to assess the retrieval performances in CBIR, various indices can be used (see [43] for a review). Between these, the most commonly used indices are Precision and Recall, usually presented as a Precision versus Recall graph (PR graph). The Precision corresponds to the average fraction of retrieved patches that are relevant amongst the n top closest patches, whereas the Recall is the average fraction of relevant patches that are retrieved in the top n closest image. The PR graph is constructed by varying the parameter n between 1 and N_T, the total number of patches in the database. This is equivalent to varying the distance threshold between the query image and the other images of the database. The Recall increases with n, whereas the Precision decreases when n increases. A texture model is therefore considered more performant than others when its associated PR curve is closer to the upper-right corner of the graph. This graphic representation is similar to ROC curves which plots the true positive rate versus the true negative rate. Nonetheless, as recommended in [44], the use of PR graphs was preferred as it can be more informative when the analysed database is unbalanced in terms of number of samples per class. The PR graph gives a general insight on the ability of a model to detect all the images of a same class without having to scan the entire database. The interpretation of PR graphs is used here to identify which multivariate model would be the best choice for the classification of VHR remote sensing data. In the next part of this section, the proposed strategy for using wavelet-based multivariate models in a supervised classification is presented.

B. Supervised Classification

The classification of VHR optical data using texture features can be addressed in different manners. The most intuitive way is to proceed with a pixel-wise approach using a sliding window to characterize local spatial dependencies. This approach is often preferred when using GLCM descriptors resulting in images of parameters subsequently used as inputs in a classifier [6], [9]. This classification technique may have

however some drawbacks: (1) the optimal size of the sliding window is not easily determined automatically and trials and errors are often needed to fix it; (2) the VHR of optical data readily translates into high local variability which can be difficult to fully capture during the learning phase without requiring a large database; (3) the computational cost of a pixel-wise analysis can be prohibitive given the size of the objects to classify in regard to the spatial resolution of the image; (4) the use of a sliding window can be the source of classification errors especially at the vicinity of object edges where mixed textures may be present. For all these reasons, we opted for a region-wise classification based on a pre-segmentation of the image to classify as advocated for example in [45], [46]. This alternative enables to reduce the number of entities to classify to the number of regions obtained by the segmentation and thereby to reduce the computational cost and to increase the homogeneity of the classification results. The proposed supervised classification method operates as follows.

During the learning phase, texture features are extracted from each patch of a learning texture database. This database is the same as the one used during the CBIR analysis. As a reminder, this database was created by extracting texture-wise homogeneous regions representative of the different classes of interest. In parallel, the same texture features are extracted from each region identified during the pre-segmentation. In the case of multivariate models, the wavelet decomposition is applied on the entire image to classify. At each scale of the decomposition, a label image derived from the pre-segmentation is downsampled to the size of the wavelet subbands so that the wavelet coefficients corresponding to each region can be easily identified. The multivariate modeling is next performed by only considering the neighborhoods contained in each region within each subband. Once texture features are extracted from the learning database and form the regions of the pre-segmentation, a classifier is used to determine which class is affected to each region according to its decision rules. The classification results is further validated by comparing it with a ground truth if available.

Any kind of classification algorithms can be adapted to this classification scheme as long as it includes a notion of similarity or if it respects the probabilistic nature of the proposed multivariate models. Three types of classification algorithms are tested in the following section: k-Nearest Neighbors (k-NN), maximum likelihood (ML) and Support Vector Machine (SVM). The k-NN relies only on a similarity measure between the texture features derived from the learning samples and the region to classify. The majority class among the k-NN is affected to the region. Similarly to the CBIR analysis, a GD or KLD is used to measure similarities. In the ML classifier, a log-likelihood criterion is computed between the multivariate models estimated on the learning samples and the local observations of the spatial dependency within each region. The model learned on the regions is not used in this case. A region is classified according to the class of the sample that maximizes this likelihood criterion. The advantage of the ML classification in comparison to k-NN is that it does not require to re-estimate new features from each of the region. This can be helpful in the case of small-sized regions (over-segmentation) for which the number of local observations in each region may not be sufficient to estimate a stable model. Nonetheless, the computational cost of the ML classifier can be significantly higher, especially when a large number of regions are to be classified. To use a SVM classifier, a non-linear kernel-based transformation is first applied on the data to project them in a new space where a hyperplan between classes can be defined. To do so, a Gaussian kernel K is adapted to the similarity measure between models by replacing the Euclidean distance with the KLD or the GD [47], [48]:

$$K(x, x') = \exp\left(-\frac{\text{distance}(x, x')^2}{2\sigma^2}\right) \tag{11}$$

where (x, x') is a couple of observations and σ the standard deviation of the Gaussian kernel. The value of σ is fixed here to maximise the performance of the classifier through a cross-validation procedure.

IV. Examples of Applications

In this section, we illustrate the proposed strategy with two examples of application of wavelet-based multivariate modeling for the classification of VHR optical remote sensing data. Both examples aim at the detection of one particular land cover in VHR panchromatic Pléiades data: vineyards and cultivated oyster fields.

A. Detection of Vineyards

Because of the periodic aspect of vine rows, textural analysis for the detection of vineyards is often performed through the use of frequency-based methods relying on the Fourier transform [49], [50], [51]. Among these frequency-based methods, Rabatel *et al.* [12] suggested an unsupervised segmentation approach in which each vine plot is isolated by selecting its frequency response in the Fourier domain using adapted Gabor filters. To apply frequency-based approaches in this context, a good contrast is needed between vine rows and inter-rows for the periodic structure to be visible. But, according to the Shannon-Nyquist theorem, such structure would only be observable if the spatial resolution of the image is at least twice as small as the period of the pattern. In many French wine-growing regions and in particular in the area of Bordeaux, the plant density is such that the inter-row distance can be close to 1 meter. Using VHR panchromatic data such as those produced by the Pléiades sensor (spatial resolution of 0.7 m) can therefore be problematic. Aliasing artifacts can be common in these data. This will result in a significantly altered texture in which the periodic aspect of vine rows is no longer clearly visible. In this application, we explore the potentialities of four wavelet-based multivariate models (SCM, SIRVgauss, SIRVg0, GGC) to detect vineyards despite the presence of these artifacts. The classification results obtained with these models are compared with those produced when using GLCM descriptors and the unsupervised segmentation approach proposed in [12].

To achieve this, we used four panchromatic images acquired by the Pléiades sensor on three wine appellations within the wine-growing region of Bordeaux: Sauternes, Saint-Émilion

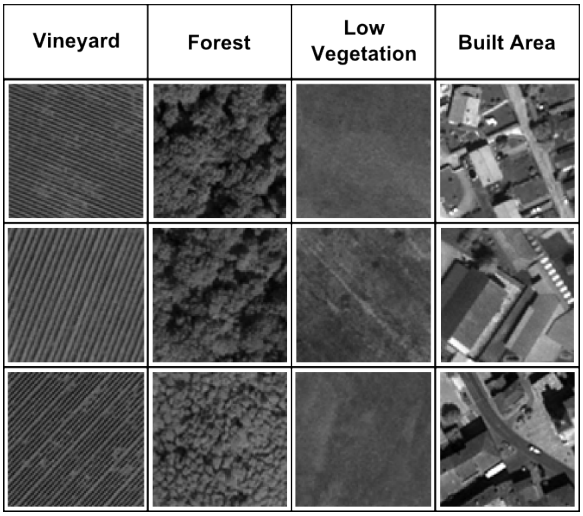


Fig. 3: Examples of texture patches used in a learning database for vineyards classification.

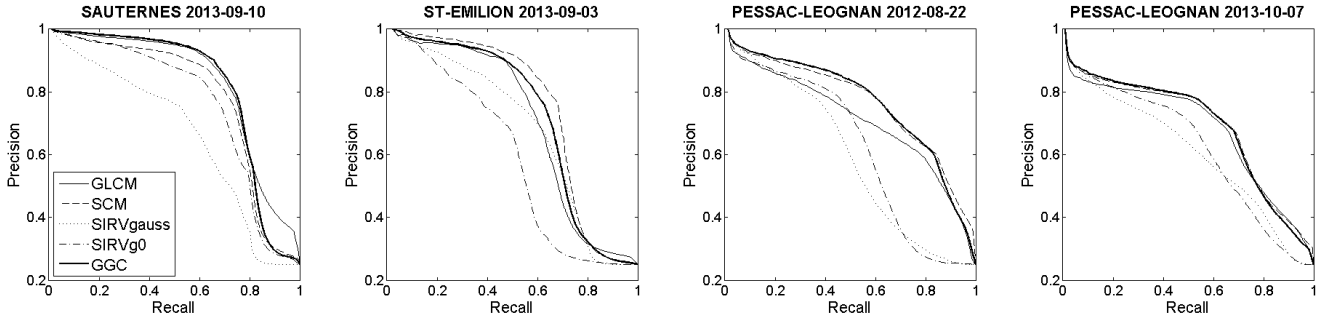
and Pessac-Léognan. Plant densities in these three appellations are such that the inter-row distances (IRD) are of approximately 1 to 1.2 m in Pessac-Léognan and 1.4 to 2 m in Saint-Émilion and Sauternes. Besides, if the spatial resolution of panchromatic Pléiades data is of 0.7 m, the Ground Sampling Distance (GSD) at the time of the image acquisition only reaches this resolution at the nadir of the satellite for an incidence angle close to 0°. For the Pléiades data used in this study, the incidence angle is significantly higher (from 13° to 33°) except for the Sauternes image (4°) and the GSD is hence significantly increased (from 0.75 to 0.9 m). Aliasing effects and the resulting alteration of textures is all the more important that the GSD is high and that the IRD is low.

From each of the four panchromatic data, a texture database is created by extracting image patches (128×128 pixels) from texture-wise homogeneous areas in the data. These selected patches are representative of the different land cover classes to distinguish in the landscape. In the wine-growing region of Bordeaux, the landscape is dominated by vineyards standing alongside typical peri-urban land covers. Four main classes are therefore considered: vineyards, forest, low vegetation (i.e. meadows, regenerating vineyards, bare soil, etc.) and built areas (Fig. 3). The texture databases constructed from each of the four panchromatic data contains 25 patches for the forest, low vegetation and built areas classes. For the vineyards, the lack of rotation invariance of the wavelet-based texture features has to be taken into consideration and a fairly exhaustive set of patches in terms of orientation has to be created. To do so, we included in the database 25 image patches of vineyards for each main classes of orientation defined by intervals of 20° between 0° and 180° (9 classes of orientation). Each database contains a total of 300 image patches with 225 for vineyards only.

The created databases are first exploited in a CBIR analysis to compare the retrieval performances of the tested multivariate models with those obtained with GLCM descriptors. In order to extract GLCM descriptors from image patches, the distance between pairs of pixels is set to 1 and four Haralick descriptors are selected: entropy, homogeneity, correlation and mean. These four descriptors are extracted from each GLCM computed in four orientations (0°, 45°, 90°, 135°) and further averaged to obtain rotation invariant descriptors [52]. These parameters are selected as they maximize the retrieval performances of the GLCM. Besides, multivariate models are applied on subbands obtained with a DWT with Daubechies db4 wavelet filters. Wavelet db4 filters are chosen as they enable a compromise between the spatial extent of the filter and its localization in frequency. The number of scales in the decomposition is set to 2 producing six wavelet subbands (2 scales × 3 orientations). Neighborhoods of 3×3 are considered within each subband to represent the observations of the local spatial dependency. A total of six sets of model parameters are extracted for each image patch of the database (one set for each wavelet subband).

The PR curves of the wavelet-based multivariate models and the GLCM descriptors are shown in Fig. 4 for each of the texture databases derived from the four Pléiades panchromatic data. It appears clearly on these curves that the retrieval performances of the multivariate models depend directly on the ratio GSD/IRD associated with the processed image data. The more this ratio increases, the more the PR curves struggle to get closer to the upper right corner of the graph, synonym of a perfect retrieval behavior. Indeed, when this ratio is lower than 0.5 (Fig. 4a), i.e. when the conditions of the Shannon-Nyquist theorem are respected, the texture associated with vineyards is less altered and retrieval performances are high with PR curves close to the upper right corner of the graph. In this case, GLCM descriptors as well as models like SCM and GGC are the more performant. Conversely, when the ratio is close to 0.5 (Fig. 4b) or slightly higher than 0.5 (Fig. 4c), the texture is potentially more altered and retrieval performances decrease. In those cases, multivariate models such as SCM and GGC display higher performances than GLCM descriptors. Finally, when the ratio GSD/IRD is significantly higher than 0.5 (Fig. 4d), the level of alteration of textures is such that none of the tested texture analysis succeeds to retrieve the textures efficiently. This lower sensitivity of wavelet-based multivariate models to the texture alteration confirms their ability to capture more informative components of the texture in the wavelet domain than GLCM descriptors in the image domain. This is in agreement with previous results [53] in which wavelet-based features were shown to be more adapted to capture lower and mid-frequency texture information, while the GLCM is more efficient to describe high-frequency textural components. Two multivariate models display significantly better retrieval performances than the two others in this particular context of application, i.e. the SCM and the GGC models. These two models are further tested in the proposed supervised classification strategy.

The texture patches database used previously in the CBIR analysis are further used in the classification process as a training database. Prior to the classification, a pre-segmentation is operated on the image using a Mean-Shift algorithm [54]. As this algorithm relies directly on the distribution of gray levels, a smoothing filter is first applied on the image to prevent



(a) GSD/IRD = 0.35 (b) GSD/IRD = 0.5 (c) GSD/IRD = 0.61 (d) GSD/IRD = 0.75

Fig. 4: Recall/Precision curves for GLCM and four multivariate models obtained with vineyard texture databases created from four panchromatic Pléiades data.

Method	OA	kappa	PA	UA	Time [1]
Gabor	87.83	0.69	68.14	89.67	23 sec.
GLCM - SVM	85.29	0.66	77.56	75.17	2 min.
SCM - ML	94.93	0.88	97.18	87.75	79 min.
GGC - ML	94.09	0.87	97.51	85.42	84 min.

OA = Overall Accuracy - PA = Producer's Accuracy - UA = User's Accuracy

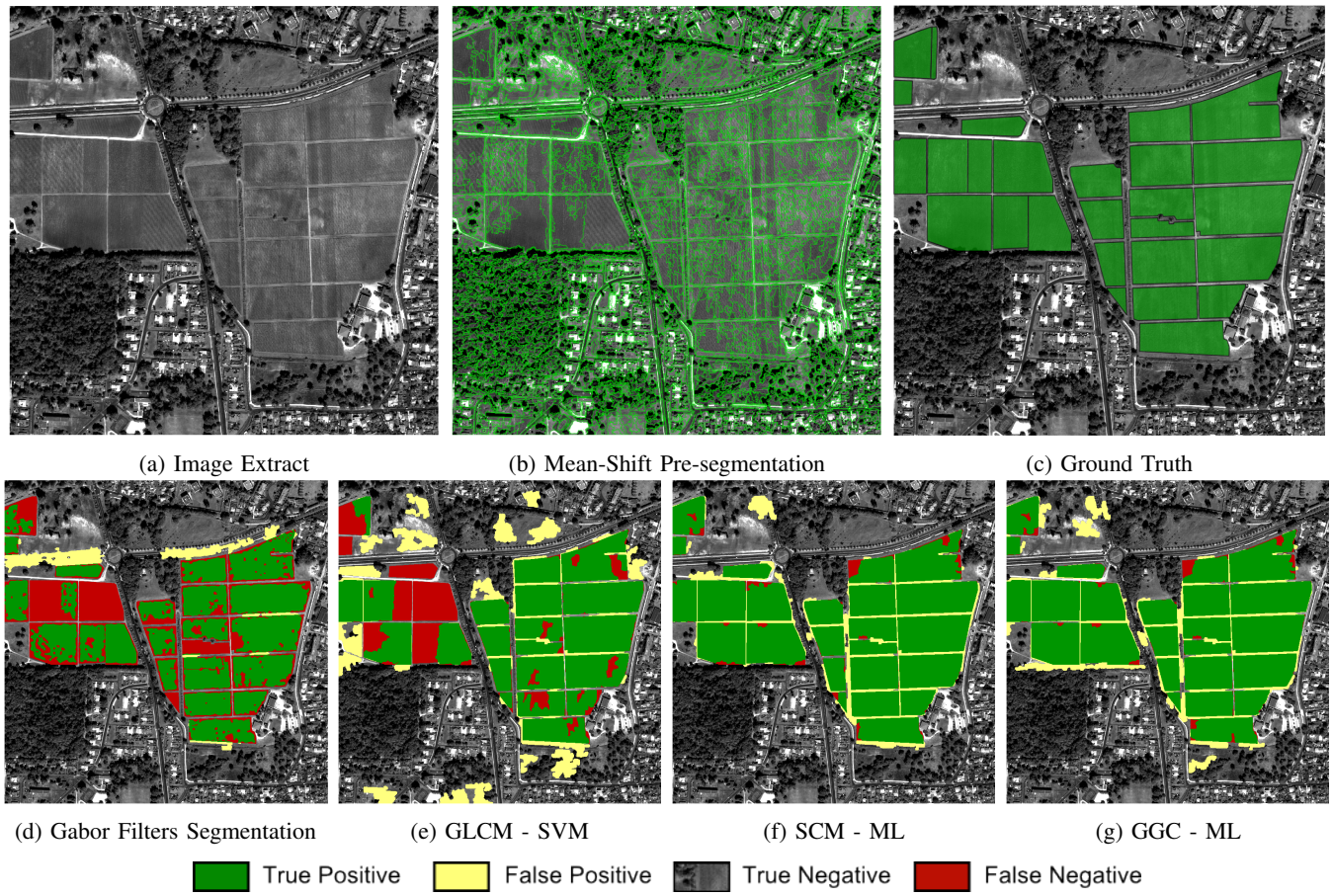
TABLE I: Accuracy indices and computational time for vineyards classification

regions from following vine rows, creating thereby irregular narrow contours. The minimum region size parameter of the Mean-Shift algorithm is also set at a low value (500 pixels) so that regions containing mixed textures are limited. The result can be therefore over-segmented. Note that other segmentation algorithms could be used for this step. The identified regions are then classified according to the chosen classifier by either extracting new features from these regions when using k-NN and SVM or by estimating a likelihood criterion when using the maximum likelihood (ML) classifier. Classification results are then compared with a ground truth for validation.

The classification results obtained with GLCM descriptors as well as with the SCM and GGC wavelet-based multivariate models are presented in Fig. 5 for an image extract of one of the Pléiades data acquired on the Pessac-Léognan appellation with a GSD/IRD ratio equal to 0.61 (Fig. 5a). These results are compared with the segmentation result achieved with the unsupervised method based on the use of Gabor filters [12]. Regions classified or segmented as vineyards are displayed in accordance with their consistency with the ground truth (Fig. 5c) by describing them in terms of true positives, false positives, true negatives and false negatives. Only the results obtained with the most efficient classifier are presented for each textural analysis appoach in Fig. 5. The corresponding accuracy indices as well as computational time are shown in Table I. These results are discussed here below.

[1] These durations were obtained with an Intel Core i5-3570 CPU at 3.40 GHz and 8 GB of RAM, running the 64 bit version of Windows 7 operating system. The codes for the supervised classification (SVM and ML) and the Gabor filters approach were implemented and run in MATLAB (version 8.5, R2015a, 64 bit). These durations do not include the Mean-Shift pre-segmentation and the model training, these steps being identical for all texture-based approaches.

Methods using texture features (GLCM and multivariate models) reveal higher detection performances than the segmentation method using Gabor filters. This lesser ability of Gabor filters to detect vine plots in this context is mainly due to the short IRD (1 to 1.2 m) which significantly contributes to the attenuation of the frequency response of the vine plots. Plots with the lowest IRD are practically undetected by this approach (see false negatives in Fig. 5d). Besides, the presence of missing vines in most of the plots interrupts locally the frequency patterns. This translates in small false negative areas appearing in the detected plots. Nevertheless, despite this under-detection problem of Gabor filters, the confusion with other land covers (false positives and user's accuracy) is low in comparison to the use of texture features. Confusions however occur with the neighboring road network which also produces a frequency response in the Fourier spectrum. Concerning the textural approaches, higher detection performances are observed when using multivariate models. Results obtained with these models (Fig. 5f-5g) display indeed higher true positive and lower false positive rates than those obtained with GLCM descriptors (Fig. 5e). This confirms the higher sensitivity to texture attenuation observed for the GLCM in the CBIR analysis. The best compromise between higher true positive and lower false positive is achieved when using the SCM model (Fig. 5f). Most of the false positives are due to confusions with low vegetation which can share similarities with vineyards especially when the texture pattern associated with vine rows is significantly altered. Also, regions defined by the pre-segmentation algorithm tend to spread beyond the limits of vine plots and create mixed regions, especially around narrow paths delimiting the plots. These detection errors are due to the Mean-Shift algorithm rather than to the texture analysis. Overall, wavelet-based multivariate models and the proposed strategy to apply them in this particular context of vineyards classification proved to produce higher detection performances than standard approaches such as GLCM descriptors or Gabor filters. These observations are confirmed by the accuracy evaluation (Table I) with higher accuracy indicators for multivariate models than for GLCM descriptors or Gabor filters. It can however be noted that the computational duration of the classification is, as expected,



(a) Image Extract (b) Mean-Shift Pre-segmentation (c) Ground Truth

(d) Gabor Filters Segmentation (e) GLCM - SVM (f) SCM - ML (g) GGC - ML

Fig. 5: Classification results on an extract of the 2012-08-22 Pléiades data acquired in Pessac-Léognan appellation. Green contours in Fig. 5b correspond to the frontiers of the regions obtained with the Mean-Shift algorithm.

significantly higher for wavelet-based mulvariate modeling with a computational cost up to 40 times higher than the use of GLCM descriptors and up to 200 times higher than the use of Gabor filters. This higher computational time is mainly due to the calculation of the maximum likelihood criterion between the multivariate models estimated on each training sample and the local observations within each region to classify.

In the following application, the proposed classification strategy is adapted to the detection of small and narrow structures in the case of cultivated oyster fields mapping.

B. Detection of Cultivated Oyster Fields

In the Arcachon Bay, a mesotidal lagoon located in the South-West of France, oysters are traditionally grown on racks of a few tens of meters to more than hundred meters long. These racks are set directly on the sediment in groups of two to five spaced by 1 to 2 meters. These groups of parallel structures are directly visible in VHR optical data and can be detected by using texture analysis in order to assess the evolution of the area occupied by this shell farming industry. Some recent studies demonstrate the interest of the international community to detect and map oyster habitat in tidal flats [55], [56]. In these studies, oyster reefs are mainly detected by making use of the radar backscattering properties specific to these land covers in SAR remote sensing data. Here, we adapt the proposed classification strategy using wavelet-based multivariate models to detect such structures in VHR optical data.

For this application, we used a panchromatic image acquired by the VHR Pléiades sensor in April 2013. This image covers a large central part of the Arcachon Bay where the oyster farming activity is predominant. The learning database used for this application is created differently than in the previous application. The use of same-sized image patches to learn texture features may indeed be less appropriate in the case of oyster fields mapping as narrow groups of oyster racks may be separated by a few meters large spaces of bare sediment. The textural signature derived from such patches would therefore include textural informations relative to both oyster racks and bare sediment. Instead, unevenly-sized Regions Of Interest (ROI) whose contours are adapted to the shapes of the structures to detect are extracted from the image. Three classes of land covers are considered: cultivated oyster fields (40 ROIs), abandoned fields (20 ROIs) and foreshores (20 ROIs) (Fig. 6). The latter corresponds to any area of bare sediment where no oyster is grown. For this class, image patches of the same size are extracted since the texture associated with foreshores is homogeneous in vast areas.

First, the learning database is analysed in a CBIR framework to compare retrieval performances and identify which multi-

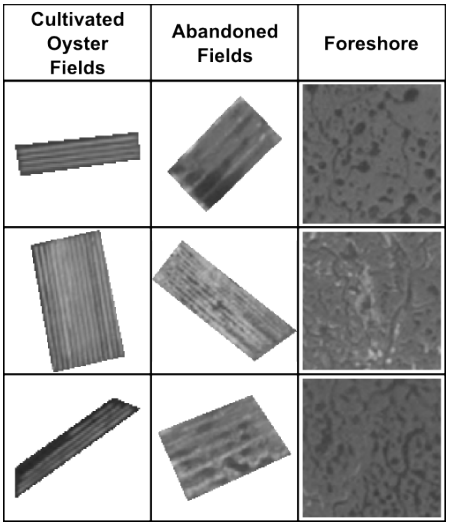


Fig. 6: Examples of texture ROI used in the learning database for oyster fields classification.

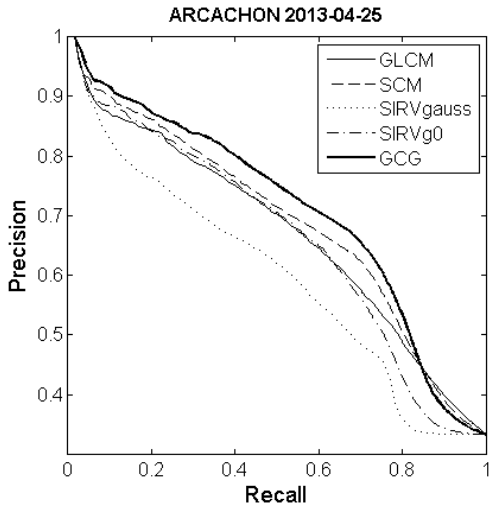


Fig. 7: Recall/Precision curve for GLCM and four multivariate models obtained with the texture database created from 2013-04-25 panchromatic Pléiades data.

variate model would perform best in the context of application. To extract GLCM texture descriptors from ROIs, the same parametrization of the GLCM than in the previous application is used, i.e. the distance between pairs of pixels is set to 1 and four Haralick descriptors (entropy, homogeneity, correlation and mean) are computed in four orientations before being averaged. Besides, multivariate modeling is applied on the six subbands obtained by applying a DWT with Daubechies db4 wavelet filters and 2 scales and 3 orientations of decomposition. Neighborhoods of 3×3 are considered within each subband to represent the observations of the local spatial dependency. As in the previous application, a total of six sets of model parameters are extracted for each image patch of the database (one set for each wavelet subband).

The computed PR curves show higher retrieval performances for almost all the multivariate models in comparison to GLCM descriptors with the exception of the SIRVgauss model (Fig. 7). It is worth noting that conversely to the application dedicated to the vineyards detection, PR curves are decreasing faster and are more distant from the upper right corner. This gives hints on the potentially higher level of confusion between the considered classes. This could be explained by the shared anisotropy between the cultivated oyster fields and abandoned fields which could be the source of classification errors. The performances of the three best multivariate models (SCM, SIRVg0 and GGC) are further assessed in classification.

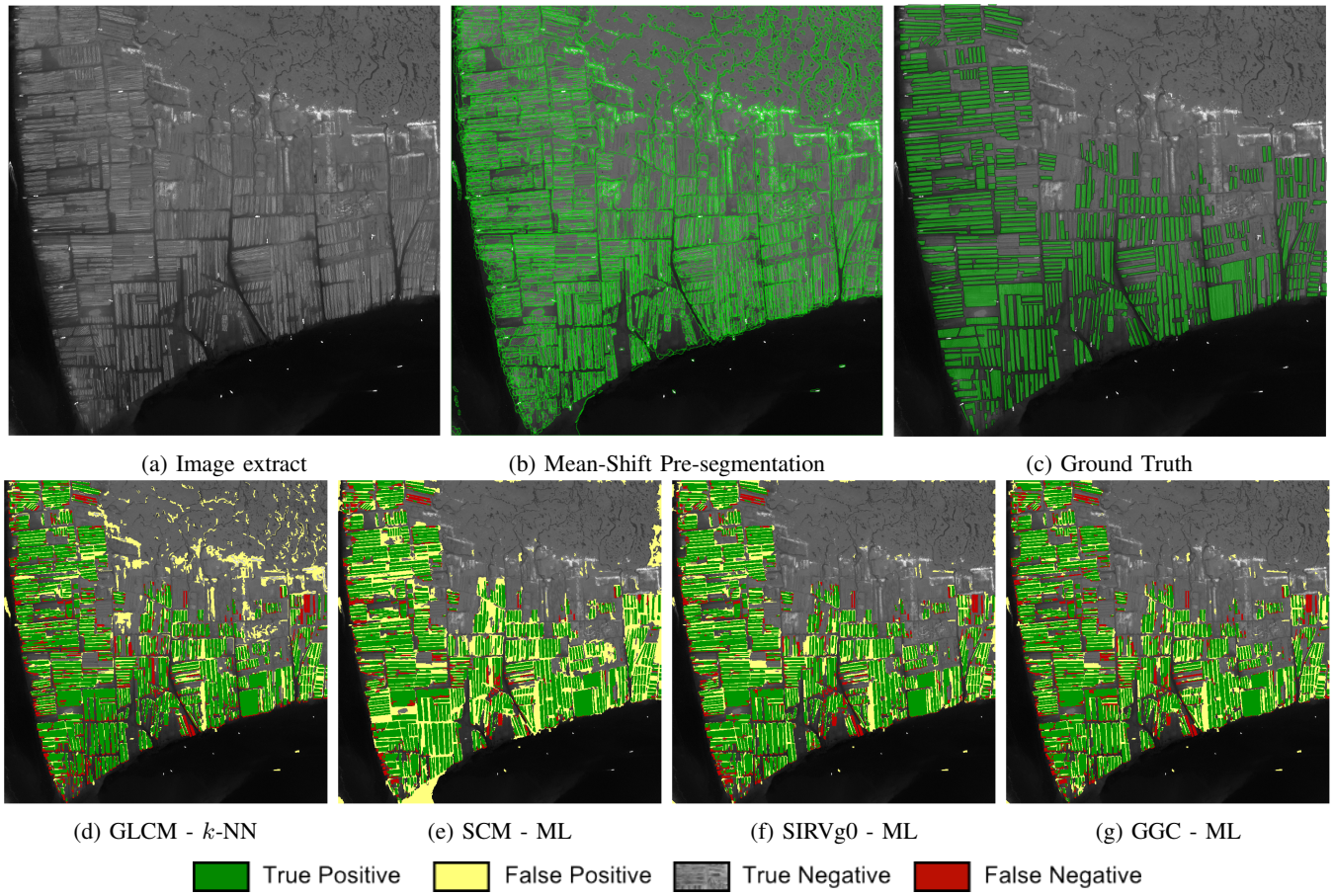
The ROI database used previously in the CBIR analysis is further used in the classification process as a training set. Prior to the classification, a Mean-Shift pre-segmentation is performed on the image to classify. As groups of oyster-growing racks can be narrow, the Mean-Shift algorithm is deliberately parametrized to create an over-segmentation with the minimum region size parameter set to 100 pixels. Small regions are hence produced (Fig. 8b). The identified regions are then classified according to the chosen classifier by either extracting new features from these regions when using k-NN classifier for GLCM descriptors or by estimating a likelihood criterion when using the maximum likelihood (ML) classifier for multivariate models. Classification results are presented in Fig. 8 for an image extract of one of the Pléiades data. As previously, regions classified as cultivated oyster fields are displayed in accordance with their consistency with the ground truth (Fig. 8c) by describing them in terms of true positives, false positives, true negatives and false negatives. Only the results obtained with the most efficient classifier are presented for each textural analysis appoach in Fig. 8. The corresponding accuracy indices and computational time results are shown in Table II. These results are discussed here below.

Method	OA	kappa	PA	UA	Time [1]
GLCM - kNN	86.28	0.66	83.13	69.2	47 sec.
SCM - ML	89.39	0.61	82.88	61.23	98 min.
SIRVg0 - ML	86.3	0.7	87.72	71.44	104 min.
GGC - ML	88.01	0.7	84.27	72.87	109 min.

OA = Overall Accuracy - PA = Producer's Accuracy - UA = User's Accuracy

TABLE II: Accuracy evaluation for oyster racks classification

The produced classification results show an equivalent capacity of two of the selected multivariate models (SIRVg0, GGC) to correctly detect oyster racks with a low proportion of false positives (Fig. 8f-8g). Conversely, the result obtained with the SCM model (Fig. 8e) reveals higher confusions with the bare sediment spaces between oyster racks with significantly higher false positives in those areas. However, as the two other multivariate models (SIRVg0 and GGC), it prevents confusion errors with abandoned fields. On the other hand, even though it achieves a satisfactory delimitation of cultivated oyster fields without significant confusion with bare sediments, the use of GLCM descriptors also induces classification errors with abandoned fields as it can be seen in the upper right part of Fig. 8d. As it was observed for the detection of vineyards, the best compromise between high level of true positives and low level of false positives is observed in this application when



(a) Image extract (b) Mean-Shift Pre-segmentation (c) Ground Truth

(d) GLCM - k-NN (e) SCM - ML (f) SIRVg0 - ML (g) GGC - ML

Fig. 8: Classification results on an extract of the 2013-04-25 Pléiades data acquired on the Arcachon Bay. Green contours in Fig. 8b correspond to the frontiers of the regions obtained with the Mean-Shift algorithm.

using wavelet-based multivariate models. This also confirms the better performances achieved with these models in the CBIR analysis in comparison to GLCM descriptors. These observations are confirmed by the accuracy evaluation (Table I) with higher accuracy indicators for the SIRVg0 and GGC models than for SCM and GLCM descriptors. As noted previously for the detection of vineyards, the computational cost of the maximum likelihood classification with waveled-based multivariate models is as expected significantly higher than the one observed for the k-NN classifier with GLCM descriptors.

C. Discussion

In both applications detailed above, at least one of the tested wavelet-based multivariate models revealed higher classification accuracies than the use of GLCM descriptors, a standard approach for textural analysis in VHR optical remote sensing data. It was however observed that the models achieving the highest accuracy can be different depending on the context of application. The SCM and GGC models were indeed more accurate to detect vineyards while the SIRVg0 and GGC models displayed better performances in the case of oyster fields mapping. In a recent work [27], a similar study was carried on the classification of mono-specific forest stands and showed that the SIRVgauss model was the best choice for this particular application. Each context of application has indeed its own specificities in terms of size, frequency and anisotropy of the textures to differentiate. The best model to be used can thereby be different depending on these specificities. A CBIR analysis carried on the learning database has proven to be a good way to identify the most adequate model prior to the classification.

Overall, this confirms the interest of the use of multivariate models to characterize textures and the potentially richer textural information captured through this modeling approach in the wavelet domain. The proposed supervised classification strategy relying on the use of a learning database composed of texture-wise homogeneous patches or ROI and a pre-segmentation of the image to classify also revealed to be an appropriate procedure and achieved satisfactory classification results in both applications. This strategy was applied by using a DWT, this choice being motivated by the ability of the DWT to produce decorrelated subbands. Nonetheless, any other types of wavelet decompositions (stationary wavelet transform, complex wavelet transform, steerable pyramids, etc.) could be used. Besides, one major quality of the proposed strategy is that only a few parameters need to be set. The only choices to be made concern the type of wavelet decomposition and the number of scales of this decomposition, as well as the eventual parameters related to the segmentation algorithm.

Once these options are set, the rest of the procedure is straightforward and can be easily adapted to any kind of texture-based classification problems.

It is also worth noting that texture features derived from DWT wavelet subbands are specific to the orientation of each subband and features are not in essence rotation invariant. Therefore, two identical anisotropic textures only differing by their main orientation could be considered as different or dissimilar when using such features. This problem comes from the similarity measure used between models as the distance is measured between pairs of orientation specific subbands. In the applications exposed above, the lack of rotation invariance was compensated by using a fairly exhaustive learning database in which all possible orientations of the textures are represented. In other cases, this exhaustivity can be more difficult to reach. In [27], a solution was proposed by applying a rotation to each region of the learning database and to each region to classify in order to force the main orientation to 0° before proceeding with the feature extraction. This solution proved to be efficient with a significant increase in classification accuracy.

Finally, we employed here a Mean-Shift algorithm to perform the pre-segmentation to identify homogeneous regions in the image to classify. As it was observed in the two presented applications, this algorithm has some drawbacks, especially in narrow areas where regions of mixed textures can be created. A pre-treatment filtering process could be used prior to the segmentation to avoid these mixed texture regions by applying for example anisotropic diffusion filters [57] or by using morphological neighborhoods [58]. Besides, any other types of segmentation algorithm could be used to produce a pre-partition of the image.

V. Conclusion

Although wavelet-based probabilistic models gained in popularity in the image processing and computer vision communities in the recent years, they have not yet been extensively explored for applications in VHR optical remote sensing data. Here, we gathered evidence that this approach is also perfectly suitable for the supervised classification of land covers in such VHR optical data. Moreover, we proposed a complete strategy to apply wavelet-based multivariate models in a supervised classification procedure of textures in VHR panchromatic data. This strategy relies on a learning texture database made of texture patches or ROI and on a pre-partition of the image to classify. Texture features are extracted from the learning database and from the regions of the pre-partition by using multivariate models (SCM, SIRVgauss, SIRVg0, GGC) to represent the distribution of observed local spatial dependencies in wavelet subbands in a multi-scale and multi-orientation framework. A CBIR analysis carried on the learning database is first conducted to identify the most efficient models to retrieve textures in the context of application. A classifier based on a similarity measure or a likelihood criterion is next used to produce classification results with the most performant models. The applicability of this strategy was tested in two distinct contexts. In both applications, the use of the proposed strategy has enabled to achieve satisfactory classification results with at least one of the tested multivariate models displaying higher classification accuracies than the standard texture analysis using GLCM descriptors. These results confirm the pertinence of using multivariate modeling in the wavelet domain to capture potentially complex texture patterns and improve the classification of VHR optical remote sensing data. Moreover, the classification is almost straightforward with only a few parameters to be set. In terms of perspective, this strategy and the use of multivariate models could also be tested in a context of multi-sensor data fusion by combining the modeling of dependencies between color bands in mutispectral data or polarimetric bands in SAR data with the modeling of spatial dependencies in panchromatic image.

Acknowledgment

The authors would like to thank the Recette Thématique Utilisateur/ORFEO and ISIS program of the CNES for providing the Pléiades images used in this study.

References

[1] R. M. Haralick, K. Shanmugam, and I. K. Dinstein, "Textural features for image classification," *IEEE Transactions on Systems, Man and Cybernetics*, vol. SMC-3, no. 6, pp. 610–621, 1973.

[2] F. Pacifici, M. Chini, and W. J. Emery, "A neural network approach using multi-scale textural metrics from very high-resolution panchromatic imagery for urban land-use classification," *Remote Sensing of Environment*, vol. 113, no. 6, pp. 1276–1292, 2009.

[3] M. Pesaresi, A. Gerhardinger, and F. Kayitakire, "A robust built-up area presence index by anisotropic rotation-invariant textural measure," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 1, no. 3, pp. 180–192, 2008.

[4] C. A. Coburn and A. C. B. Roberts, "A multiscale texture analysis procedure for improved forest stand classification," *International journal of remote sensing*, vol. 25, no. 20, pp. 4287–4308, 2004.

[5] S. E. Franklin, R. J. Hall, L. M. Moskal, A. J. Maudie, and M. B. Lavigne, "Incorporating texture into classification of forest species composition from airborne multispectral images," *International Journal of Remote Sensing*, vol. 21, no. 1, pp. 61–79, 2000.

[6] B. Béguet, S. Boukir, D. Guyon, and N. Chehata, "Modelling-based feature selection for classification of forest structure using very high resolution multispectral imagery," in *2013 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE, 2013, pp. 4294–4299.

[7] I. Champion, C. Germain, J. P. Da Costa, A. Alborini, and P. Dubois-Fernandez, "Retrieval of forest stand age from SAR image texture for varying distance and orientation values of the gray level co-occurrence matrix," *IEEE Geoscience and Remote Sensing Letters*, vol. 11, no. 1, pp. 5–9, 2014.

[8] F. Kayitakire, C. Hamel, and P. Defourny, "Retrieving forest structure variables based on image texture analysis and IKONOS-2 imagery," *Remote Sensing of Environment*, vol. 102, no. 3, pp. 390–401, 2006.

[9] T. A. Warner and K. Steinmaus, "Spatial classification of orchards and vineyards with high spatial resolution panchromatic imagery," *Photogrammetric Engineering & Remote Sensing*, vol. 71, no. 2, pp. 179–187, 2005.

[10] A. Balaguer, L. A. Ruiz, T. Hermosilla, and J. A. Recio, "Definition of a comprehensive set of texture semivariogram features and their evaluation for object-oriented image classification," *Computers & Geosciences*, vol. 36, no. 2, pp. 231–240, 2010.

[11] S. Aksoy, H. G. Akcay, and T. Wassenaar, "Automatic mapping of linear woody vegetation features in agricultural landscapes using very high resolution imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 1, pp. 511–522, 2010.

[12] G. Rabatel, C. Delenne, and M. Deshayes, "A non-supervised approach using Gabor filters for vine-plot detection in aerial images," *Computers and Electronics in Agriculture*, vol. 62, no. 2, pp. 159–168, 2008.

[13] A. Lucieer and H. van der Werff, "Panchromatic wavelet texture features fused with multispectral bands for improved classification of high-resolution satellite imagery," in *2007 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*. IEEE, 2007, pp. 5154–5157.

[14] L. A. Ruiz, A. Fdez-Sarría, and J. A. Recio, "Texture feature extraction for classification of remote sensing data using wavelet decomposition: a comparative study," in *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2004, vol. XXXV-B4, pp. 1109–1114.

[15] Q. Jackson and D. A. Landgrebe, "Adaptive bayesian contextual classification based on Markov random fields," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 40, no. 11, pp. 2454–2463, 2002.

[16] Y. Zhao, L. Zhang, P. Li, and B. Huang, "Classification of high spatial resolution imagery using improved Gaussian Markov random-field-based texture features," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, no. 5, pp. 1458–1468, May 2007.

[17] C. Vaduva, I. Gavat, and M. Datcu, "Deep learning in very high resolution remote sensing image information mining communication concept," in *Signal Processing Conference (EUSIPCO), 2012 Proceedings of the 20th European*. IEEE, 2012, pp. 2506–2510.

[18] A. Romero, C. Gatta, and G. Camps-Valls, "Unsupervised deep feature extraction for remote sensing image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. PP, no. 99, pp. 1–14, 2015.

[19] M. N. Do and M. Vetterli, "Wavelet-based texture retrieval using generalized Gaussian density and Kullback-Leibler distance," *IEEE Transactions on Image Processing*, vol. 11, no. 2, pp. 146–158, 2002.

[20] S. G. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 7, pp. 674–693, 1989.

[21] S. K. Choy and C. S. Tong, "Statistical wavelet subband characterization based on generalized gamma density and its application in texture retrieval," *IEEE Transactions on Image Processing*, vol. 19, no. 2, pp. 281–289, 2010.

[22] G. Verdoolaege, S. De Backer, and P. Scheunders, "Multiscale colour texture retrieval using the geodesic distance between multivariate generalized Gaussian models," in *2008 IEEE International Conference on Image Processing (ICIP)*. IEEE, 2008, pp. 169–172.

[23] L. Bombrun, S. N. Anfinsen, and O. Harant, "A complete coverage of log-cumulant space in terms of distributions for polarimetric SAR data," *2011 International Workshop on Science and Applications of SAR Polarimetry and Polarimetric Interferometry (POLinSAR)*, p. 136, 2011.

[24] R. Kwitt and A. Uhl, "A joint model of complex wavelet coefficients for texture retrieval," in *2009 IEEE International Conference on Image Processing (ICIP)*. IEEE, 2009, pp. 1877–1880.

[25] R. Pickard, C. Graszyk, S. Mann, J. Wachman, L. Pickard, and L. Campbell, "Vistex database," *Media Lab., MIT, Cambridge, Massachusetts*, 1995.

[26] P. Brodatz, *Textures: a photographic album for artists and designers*, vol. 66, Dover New York, 1966.

[27] O. Regniers, L. Bombrun, D. Guyon, J. C. Samalens, and C. Germain, "Wavelet-based texture features for the classification of age classes in a maritime pine forest," *Geoscience and Remote Sensing Letters, IEEE*, vol. 12, no. 3, pp. 621–625, 2015.

[28] G. Camps, D. Tuia, L. Gomez, S. Jiménez, and J. Malo, "Chapter 2 - the statistics of remote sensing images," in *Remote Sensing Image Processing*. Morgan & Claypool, 2011.

[29] I. W. Selesnick, R. G. Baraniuk, and N. C. Kingsbury, "The dual-tree complex wavelet transform," *IEEE Signal Processing Magazine*, vol. 22, no. 6, pp. 123–151, 2005.

[30] A. Laine and J. Fan, "Texture classification by wavelet packet signatures," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 15, no. 11, pp. 1186–1191, 1993.

[31] M. Unser, "Texture classification and segmentation using wavelet frames," *IEEE Transactions on Image Processing*, vol. 4, no. 11, pp. 1549–1560, 1995.

[32] E. P. Simoncelli and W. T. Freeman, "The steerable pyramid: A flexible architecture for multi-scale derivative computation," in *1995 IEEE International Conference on Image Processing (ICIP)*. IEEE Computer Society, 1995, vol. 3, pp. 3444–3444.

[33] F. Pascal, P. Forster, J. P. Ovarlez, and P. Larzabal, "Performance analysis of covariance matrix estimates in impulsive noise," *IEEE Transactions on Signal Processing*, vol. 56, no. 6, pp. 2206–2217, 2008.

[34] C. C. Freitas, A. C. Frery, and A. H. Correia, "The polarimetric $\mathcal{G}$ distribution for SAR data analysis," *Environmetrics*, vol. 16, no. 1, pp. 13–31, 2005.

[35] G. Vasile, J.-P. Ovarlez, F. Pascal, and C. Tison, "Coherency matrix estimation of heterogeneous clutter in high-resolution polarimetric sar images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 4, pp. 1809–1826, 2010.

[36] F. Pascal, Y. Chitour, J. P. Ovarlez, P. Forster, and P. Larzabal, "Covariance structure maximum-likelihood estimates in compound gaussian noise: Existence and algorithm analysis," *IEEE Transactions on Signal Processing*, vol. 56, no. 1, pp. 34–48, 2008.

[37] Y. Stitou, N. E. Lasmar, and Y. Berthoumieu, "Copulas based multivariate gamma modeling for texture classification," in *2009 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2009, pp. 1045–1048.

[38] N. E. Lasmar and Y. Berthoumieu, "Gaussian copula multivariate modeling for texture image retrieval using wavelet transforms," *IEEE Transactions on Image Processing*, vol. 23, no. 5, pp. 2246–2261, May 2014.

[39] A. Sklar, "Random variables, joint distribution functions, and copulas," *Kybernetika*, vol. 9, no. 6, pp. 449–460, 1973.

[40] A. W. M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, "Content-based image retrieval at the end of the early years," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 12, pp. 1349–1380, 2000.

[41] L. Bombrun, Y. Berthoumieu, N. E. Lasmar, and G. Verdoolaege, "Multivariate texture retrieval using the geodesic distance between elliptically distributed random variables," in *2011 IEEE International Conference on Image Processing (ICIP)*. IEEE, 2011, pp. 3637–3640.

[42] C. Lenglet, M. Rousson, R. Deriche, and O. Faugeras, "Statistics on the manifold of multivariate normal distributions: Theory and application to diffusion tensor MRI processing," *Journal of Mathematical Imaging and Vision*, vol. 25, no. 3, pp. 423–444, 2006.

[43] H. Müller, W. Müller, D. McG. Squire, S. Marchand-Maillet, and T. Pun, "Performance evaluation in content-based image retrieval: overview and proposals," *Pattern Recognition Letters*, vol. 22, no. 5, pp. 593–601, 2001.

[44] J. Davis and M. Goadrich, "The relationship between Precision-Recall and ROC curves," in *Proceedings of the 23rd international conference on Machine learning*. ACM, 2006, pp. 233–240.

[45] L. Gueguen, "Classifying compound structures in satellite images: A compressed representation for fast queries," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 4, pp. 1803–1818, 2015.

[46] L. Fang, S. Li, X. Kang, and J. A. Benediktsson, "Spectral–spatial classification of hyperspectral images with a superpixel-based discriminative sparse model," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 8, pp. 4186–4201, 2015.

[47] N. Vasconcelos, P. Ho, and P. Moreno, "The Kullback-Leibler kernel as a framework for discriminant and localized representations for visual recognition," in *2004 European Conference on Computer Vision (ECCV)*, pp. 430–441. Springer, 2004.

[48] S. Jayasumana, R. Hartley, M. Salzmann, H. Li, and M. Harandi, "Kernel methods on the riemannian manifold of symmetric positive definite matrices," in *2013 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, 2013, pp. 73–80.

[49] T. Wassenaar, J. M. Robbez-Masson, P. Andrieux, and F. Baret, "Vineyard identification and description of spatial crop structure by per-field frequency analysis," *International Journal of Remote Sensing*, vol. 23, no. 17, pp. 3311–3325, 2002.

[50] T. Ranchin, B. Naert, M. Albuisson, G. Boyer, and P. Astrand, "An automatic method for vine detection in airborne imagery using the wavelet transform and multiresolution analysis," *Photogrammetric Engineering and Remote Sensing*, vol. 67, no. 1, pp. 91–98, 2001.

[51] S. Aksoy, I. Z. Yalniz, and K. Tasdemir, "Automatic detection and segmentation of orchards using very high resolution imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 8, pp. 3117–3131, Aug 2012.

[52] R. M. Haralick, "Statistical and structural approaches to texture," *Proceedings of the IEEE*, vol. 67, no. 5, pp. 786–804, 1979.

[53] Y. O. Ouma, R. Tateishi, and J. T. Sri-Sumantyo, "Urban features recognition and extraction from very-high resolution multi-spectral satellite imagery: A micro-macro texture determination and integration framework," *IET Image Processing*, vol. 4, no. 4, pp. 235–254, 2010.

[54] D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 5, pp. 603–619, 2002.

[55] B. H. Choe, D. J. Kim, J. H. Hwang, Y. Oh, and Wooil M W. M. Moon, "Detection of oyster habitat in tidal flats using multi-frequency polarimetric SAR data," *Estuarine, Coastal and Shelf Science*, vol. 97, pp. 28–37, 2012.

[56] M. Gade, S. Melchionna, K. Stelzer, and J. Kohlus, "Multi-frequency SAR data help improving the monitoring of intertidal flats on the German North Sea coast," *Estuarine, Coastal and Shelf Science*, vol. 140, pp. 32–42, 2014.

[57] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 7, pp. 629–639, 1990.

[58] M. Fauvel, J. Chanussot, and J. A. Benediktsson, "A spatial–spectral kernel-based approach for the classification of remote sensing images," *Pattern Recognition*, vol. 45, no. 1, pp. 381–392, 2012.