A Spatial-Spectral Kernel Based Approach for the Classification of Remote Sensing Images

M. Fauvel[a,*], J. Chanussot[b], J.A. Benediktsson[c]

[a]INRA, DYNAFOR, BP 32607 - Auzeville-Tolosane 31326 - Castanet Tolosan - FRANCE
[b]GIPSA-lab, Departement Image Signal, BP 46 - 38402 Saint Martin d'Hères - FRANCE
[c]Dept. of Electrical and Computer Engineering, University of Iceland Hjardarhagi 2-6, 107 Reykjavik - ICELAND

Abstract

Classification of remotely sensed images with very high spatial resolution is investigated. The proposed method deals with the joint use of the spatial and the spectral information provided by the remote sensing images. A definition of an adaptive neighborhood system is considered. Based on morphological area filtering, the spatial information associated with each pixel is modeled as the set of connected pixels with an identical grey value (flat zone) to which the pixel belongs: The pixel's neighborhood is characterized by the vector median value of the corresponding flat zone. The spectral information is the original pixel's value, be it a scalar or a vector value. Using kernel methods, the spatial and spectral information are jointly used for the classification through a support vector machine formulation. Experiments on hyperspectral and panchromatic images are presented and show a significant increase in classification accuracies for peri-urban area: For instance, with the first data set, the overall accuracy is increased from 80% with a conventional support vectors machines classifier to 86% with the proposed approach. Comparisons with other contextual methods show that the method is competitive.

Keywords: Hyperspectral remote sensing images, urban area, adaptive neighborhood, area filtering, mathematical morphology, support vectors machines, composite kernel.

1. Introduction

The classification of optical urban remote-sensing images has become a challenging problem, due to recent advances in remote sensor technology [1]. Spatial resolution is now as high as 0.75 meter for several satellites, *e.g.* IKONOS, QUICKBIRD, and soon PLEIADES: For the same location, a panchromatic image with 0.75-meter spatial resolution and a multispectral image with 3-meter spatial resolution are available. Moreover, new hyperspectral sensors can simultaneously collect over a hundred spectral bands of an area, with increasing spatial resolution, *e.g.* 1.5 meter for airborne sensors. As a result, with such resolution many small objects and materials can now be extracted with very fine accuracy for detection, classification or segmentation. For instance, the problem of detecting or classifying *urban areas* in remotely-sensed images with lower spatial

---

*Corresponding author

Email addresses: `mathieu.fauvel@ensat.fr` (M. Fauvel), `jocelyn.chanussot@gipsa-lab.grenoble-inp.fr` (J. Chanussot), `benedikt@hi.is` (J.A. Benediktsson)

resolution has now become the even bigger problem of analyzing *within urban area structures*. Many of the applications remain to be explored and specific methodologies need to be developed to handle the complex properties of very high resolution images.

For the critical problem of land cover classification, it is conventional to use spectral information as an input to the classifier [2]. However, this entails processing a very high volume of data with high dimensionality: For instance, when dealing with hyperspectral images, each pixel-vector $\mathbf{x}$ is composed of over hundred spectral components ($\dim(\mathbf{x}) > 100$) and in such a high dimensional space, statistical estimation is a difficult task [3]. For the purpose of classification or segmentation, these problems are related to the "curse of dimensionality". As an example, the required number of training pixels for a reliable estimation is related to the square of the dimensionality for a quadratic classifier (*e.g.* the Gaussian Maximum Likelihood) [4]. However, in remote sensing applications, only limited labeled reference pixels are usually available. The consequence is the so-called "Hughes phenomenon" [5, 6]: "*With a fixed design pattern sample, recognition accuracy can first increase as the number of measurements made on a pattern increases, but decay with measurement complexity higher than some optimum value*". These problems are particularly important in classification of urban remote sensing images, where it is desirable to use the information from the spectral domain together with the information from the spatial domain. Actually, for a given pixel, it is possible to extract the size, shape, and gray-level distribution of the structure to which it belongs [7]. This information will not be the same if the pixel belongs to a roof or to a vegetation area. This is also a way to discriminate various structures made of the same materials. If spectral information alone is used, the roofs of a private house and of a larger building will be detected as the same type of structure. But using additional spatial information – the size of the roof, for instance – it is possible to classify these into two separate classes. Consequently, a contextual spatial-spectral classifier is needed to better classify urban remote-sensing images.

One major problem is the definition of a convenient multivariate statistical model that exploits both the spectral and the spatial information. Accordingly, conventional parametric statistical methods are not appropriate for the problem of combining spatial and spectral information. Consequently several methods have been proposed to analyze very high resolution remote sensing images [8]. Landgrebe and co-workers were probably the first to propose a contextual classifier, the well-known ECHO [9]. It is based on a segmentation algorithm and a region based statistical classification. A Gaussian Markov Random Field (MRF) was investigated in textural discrimination for remote sensing images segmentation in [10], involving several strategies for the estimation of MRF parameters. Later, Landgrebe and Jackson proposed an iterative statistical classifier based on MRF modeling [11]. However, MRF modeling suffers from high spatial resolution: Neighboring pixels are highly correlated and the standard neighbor system definition does not contain enough samples to be effective (here "sample" refers to pixel from the image, being a scalar or a vector). Unfortunately, a larger neighbor system entails intractable computational problems, thereby limiting the benefits of conventional MRF modeling. Furthermore, algorithms involving MRF-based strategies traditionally require an iterative optimization step, such as simulated annealing, which is extremely time consuming with high resolution data. Therefore, the use of spatial information ought to be considered with less demanding approaches in terms of computation. Benediktsson *et al.* have proposed using advanced morphological filters and a neural network based classifier as an alternative way of performing contextual classification [12]. Rather than defining a crisp neighbor set for every pixel, morphological filters enable the analysis of each pixel's neighborhood according to the structures to which it belongs. Despite good results in terms of classification accuracy, these approaches still suffer from the high dimensionality of the data, in the spectral or the spatial

domain, and need advanced pre-processing to reduce the dimensionality (*e.g.* feature extraction).

Support Vector Machines (SVM) [13] have been investigated intensively over the past few years as an alternative approach to the usual statistical and neural classifiers in high dimensional images [14]. Recently, Guo *et al.* have investigated matched kernels for the classification of hyperspectral images [15] where the parameters of the kernel are tuned adaptively by weighting each spectral band, according to their usefulness. Some approaches have been proposed to include spatial information in the SVM classification process. In [16], the authors built kernel functions that use neighborhood information. In their approach, neighbors of a pixel were defined as the pixels which belong to a square centered on the initial pixel. Then, the spatial information was modeled as the mean and variance of the gray value distribution of neighboring pixels. Another approach can be found in [17], where the spatial information was modeled as textural information: The authors proposed a wavelet-based multi-scale strategy to characterize local texture, taking the physical nature of the data into account. Then the extracted textural information was used as new feature to build a texture kernel and the final kernel was the weighted sum of a kernel made with spectral information and the texture kernel. These two previous approaches addressed the problem of merging spatial and spectral information as kernel definition problems. A new approach has been proposed in [18] where inter-pixel dependency was modeled as the mean of pixel gray values from a pixel's neighborhood system. This information was directly included in the training process as a new constraint for the optimization problem. However, only the spatial information from the support vectors is used for the final classification. Another approach using kernels on a segmentation graph was proposed by Harchaoui and Bach [19]. But their family of kernels was defined between images and not between structures in images, which significantly differs from the proposed approach.

The results achieved by the different approaches on several images demonstrate clearly the importance of a contextual spatial-spectral kernel-based classifier in the analysis of remote sensing images. A common drawback to the methods presented up to this point is the neighborhood system definition. It is based on a low-level image analysis: Starting from each pixel, the inter-pixel dependency is defined locally. For [16] and [18], the neighborhood system was a square centered on the pixel being considered, while for the wavelet transform, the neighborhood system is completely defined by the mother wavelet: Only neighbors in 4-connectivity are considered. These approaches are not appropriate for pixels located on the boundary of a structure: The fixed shape neighborhood then includes pixels from different structures. For example, as shown in Figure 4.(a), the classification of the marked pixel (roof) may be influenced by neighboring pixels actually belonging to the street. In this case, the inter-pixel dependency of a structure can be poorly estimated, leading to so-called "border effect" problems. Hence, a fixed shape or size for the neighborhood system cannot correctly handle the definition of the neighborhood system for complex images. In this article, the use of an advanced morphological filter is proposed to define adaptive neighbors based on a high-level image analysis. The idea is to define neighbors in a *structural sense*, *i.e.* to look at neighboring pixels that belong to the same structure. Note that the idea of adaptive neighbors is also investigated in non-stationary MRF, see [20] for instance and in post-classification procedure, see [21].

In this paper, the approach is to construct the neighborhood system, namely the *morphological neighborhood*, by area filtering. The objective of the approach is to combine low-level information (from the spectral domain) and high-level information (from the spatial domain). The original image is filtered with a flat zone area filter. This filter removes all the structures of the image that contain fewer pixels than a given parameter [22]. All the remaining flat zones are labeled [23]. Then the neighborhood system of a pixel is defined as the set of pixels belonging to the same flat/labeled

zone of the filtered image. Using SVM classifiers, pixels are classified by their spectrum and the statistical characteristics of their flat zone. Following [16, 17], the information is merged using a combination of kernels (or weighted sum of kernels) during the classification process. Note that the proposed neighborhood is defined pixelwise, rather than objectwise *e.g.* [24], and structures are defined as the set of pixels that share the same neighborhood.

The paper is organized as follows. In Section 2, the area filtering approach and the neighborhood system are discussed. SVM and kernel definition are presented in Section 3. Experimental results are reported in Section 4. Discussions and comparisons with other contextual methods are given in Section 5. Conclusions are drawn in Section 6.

2. Morphological Neighborhood

In this section, some basics of mathematical morphology are first reviewed. Attention is paid to image simplification using morphological filters. Then a "flat zone" area filter is presented, leading to the definition of the morphological neighborhood.

2.1. Introduction

Mathematical morphology provides high level non linear operators to analyze spatial inter-pixel dependency in an image [25]. Morphological operators have already proven their potential in remote sensing image processing [7]. Two widely used morphological operators are opening and closing by reconstructions [26]. They are connected operators that satisfy the following assertion: If the structure of the image cannot contain the structuring element (SE), then it is totally removed, else it is totally preserved. For a given SE, geodesic opening or geodesic closing provides a characterization of the size or shape of some objects present in the image: The objects that are smaller than the SE are deleted while the others (that are bigger than the SE) are preserved. To determine the shape or size of all elements present in an image, it is necessary to use a range of different SE sizes. This concept is called *Granulometry* [23, 12, 27]. When granulometry is built with connected operators, such as opening by reconstruction [26], the image is progressively simplified while no shape-noise is introduced. In that case, the resulting image contains only maxima which have a larger size than the structuring element of size λ: The structuring element can fit in each maximum. If area openings are used, the output image has its maxima that contain more than λ pixels (the area is seen as the number of pixels inside a maximum).

This concept has given rise to the Morphological Profile (MP) for the analysis of remote sensing images: The concatenation of a granulometry and anti-granulometry made with geodesic filters [28]. Figure 1 gives an example of an MP obtained with 3 openings (closings) by reconstruction with a disk, respectively, of radius 5, 13 and 21 as structuring element.

Geodesic opening and closing filters are interesting because they preserve shapes. However, they cannot provide a complete analysis of urban areas because they only act on the extrema of the image. Moreover, some structures may be darker than their neighbors in some parts of the image, yet lighter than their neighbors in others. Although this problem can be partially addressed by using an alternate sequential filter (ASF) [29], the MP thus provides an incomplete description of the inter-pixel dependency.

In [22], Soille has proposed using self-complementary filters (the definition is given in the next section) to analyze all the structures of an image, local extrema, be they minima or maxima, as well as regions with intermediate gray-levels. This assumes that any given structure of interest corresponds to one set of connected pixels. Based on an area criterion, a self-complementary flat
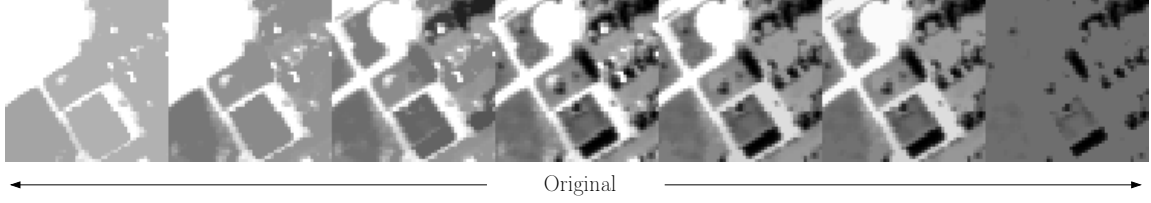
4

Figure 1: Morphological Profile: The left part of the profile corresponds to the anti-granulometry and the right part to the granulometry.

zone filter is proposed to remove small structures [22]. This kind of filter is well suited to the analysis of high resolution optical images: The very high spatial resolution results in excessively detailed data containing many irrelevant structures (*e.g.*, cars on the road). As will be detailed in the following, the area self-complementary filter is not a morphological filter, since the increasingness property no longer holds. Thus the granulometry strategy used with the MP cannot be directly applied. In this work, another approach is proposed to extract the contextual information. The idea is to build an *adaptive neighbors system for each pixel* [30], which considers neighboring pixels that belong to the same structure. In the following, the self-complementary flat zone area filter is presented as an alternative to the original granulometry operator, and the neighborhood definition is detailed.

2.2. Area Filtering

As explained in the previous section, classic opening/closing-based filters (granulometry or ASF) have the same limitation, *i.e.* they act on the maxima/minima of the image. Hence, the simplification of the image only occurs for structures that are extrema, whereas many structures corresponding to homogeneous intermediate regions are not processed. The consequence is an incomplete filtering of the structures of interest. Example of such problem is shown in Fig 2. Fortunately, flat zone approaches can tackle this problem [31].

A flat zone is a connected (in 8-connectivity) region where the gray-level is constant [32]. Flat zone filtering consists in removing all the flat zones that do not fulfill a given criterion. In this paper, the objective is to remove all the structures that are "too small" to be significant in a morphological meaning, *e.g.* the road is usually a class of interest but not the cars that might be on the road. The chosen criterion is the area of the flat zone, which is simply the number of pixels belonging to the flat zone.

Soille has proposed a flat zone filter ψ_λ^{area} based on an area criterion λ which has the following properties [22]:

- *Absorption*: The composition of two transformations ψ of different sizes λ and ν always give the result of the transformation with the biggest size parameter;

- *Self complementarity*: ψ is self-complementary with respect to the complementation operator $C \iff \psi = \psi C$.

The use of this area filter was motivated by the following peculiarities:

- The self complementarity guarantees that each structure is processed in the same way, whatever its gray value or its local contrast. Thus it analyzes all the structures of an image at
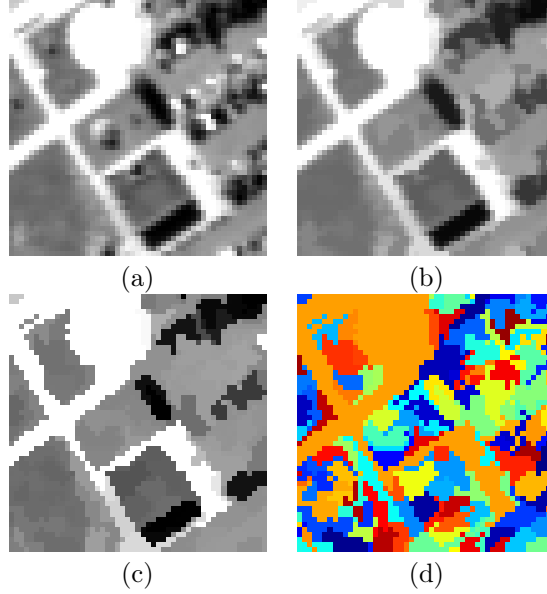
Figure 2: (a) Original image, (b) ASF based on area opening/closing, (c) flat zones area filtering and (d) the neighborhood system. For both filters λ was set to 10. Note that with ASF, many structures are of an area smaller than λ. Using ψ_λ^{area}, the number of flat zones significantly decreases, from 1995 flat zones in (a) to 127 in (c), against 1242 in (b). In (d) each color represents a set of neighbor pixels.

once, local *extrema* (be they *minima* or *maxima*) as well as regions with intermediate grey levels. Classic approaches by others workers usually have only the self dual property (*e.g.* FLST [33]).

- The algorithm can be implemented using priority queue structures and leads to very fast processing [34]. It is performed iteratively, increasing the area parameter at each step until the desired value is reached.

By using this area filter, it is possible to simplify the image by removing all structures smaller than the parameter λ. It is clear that small structures may be of interest and they can be accidentally removed when filtering. But, according to the spatial resolution of the data and the classes, it is possible to filter so as to keep only relevant structures in the image, where structures are represented as flat zones. An example of such filtering is shown in Figure 2.

2.3. Morphological Neighborhood

As stated in the introduction, the neighborhood of each pixel is defined as the connected set of pixels resulting from the application of a self-complementary area filter. This is illustrated in Figure 4 where (b) is the area filtering of (a). The filtered image is partitioned into flat zones. Each flat zone is consistent and hence belongs to one single structure in the original image. Furthermore, the smallest structures have been removed and only the main structures of interest remain. The morphological neighborhood $\Omega_{\mathbf{x}}$ of pixel $\mathbf{x}$ is defined as the set of pixels that belong to the same flat zone in the filtered image. The neighborhoods defined in this way are applied to the original image. Figure 4.(c) shows the morphological neighborhood $\Omega_{\mathbf{x}}$ associated with the observed pixel $\mathbf{x}$.

This neighborhood is obviously more homogeneous and spectrally consistent than the fixed square featured on Figure 4.(a).
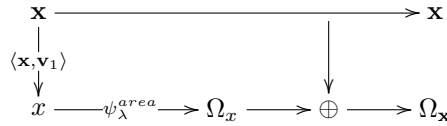
Formally, the above defined morphological neighborhood is connected to the more general concept of adaptive neighborhood in image processing [35, 36, 37, 38]. Filters based on adaptive neighborhoods are called adaptive operators: The effect of these operators is dependent of the location and of the neighborhood of the considered pixel. Such operators can be divided into two main classes: Adaptive-weighted operators and spatially-adaptive operators [35]. Morphological neighborhood belongs to the second class, the neighborhood is built adaptively for each pixel through local area and local contrast criteria. Considering another approach for defining the neighborhood adaptively, the morphological neighborhood has several interesting properties: There is no stopping criterion and there is only one parameter, the area parameter, which can be tuned using an intuitive interpretation (*e.g.*, the objects of interest must contain more than a certain amount of pixels). However, in this article we only consider crisp neighborhood: Two pixels are either neighbors or not. A more advanced approach would be to consider that two pixels are "more or less" neighbors. Such information could be included in the extraction of the spatial feature. But it goes beyond the scope of this paper.

2.4. Multichannel images

The area filter cannot be used directly on multispectral or hyperspectral remote sensing images, because of the lack of ordering relation. In order to overcome this shortcoming, several approaches can be considered, see [39] for a review of several multivariate morphological filters. Using marginal ordering, one can apply the area filter on each band independently, but considering the high inter-band correlation, this is not appropriate [40, 41]. Moreover, pixel-vectors not present in the original image can be created using marginal ordering.

Total pre-ordering was exploited in previous work. This strategy was successfully used in [42], where the Principal Component Analysis (PCA) [4] is applied to map the images onto a vector space where an actual ordering relation exists. It enables the computation of the MP on multivalued images, leading to the construction of the Extended MP (EMP) [42, 43]. Note that this EMP is somewhat different to the EMP defined in [44].

In this paper, the area filtering is computed on the first principal component to extract the neighborhood of each pixel. Then, the neighbors mask is applied on each band of the data. The diagram below sums up the methodology:

$$
\begin{array}{ccc}
\mathbf{x} & \longrightarrow & \mathbf{x} \\
\downarrow \langle \mathbf{x}, \mathbf{v}_1 \rangle & & \downarrow \\
x \xrightarrow{\psi_\lambda^{area}} \Omega_x & \longrightarrow \oplus \longrightarrow & \Omega_\mathbf{x}
\end{array}
$$

where $\mathbf{v}_1$ is the first eigenvector corresponding to the largest eigenvalue of the covariance matrix of $\mathbf{x}$ and $\oplus$ means that the one-dimensional morphological neighborhood mask is applied on each spectral band of the data. The following section details how spatial information is extracted.

2.5. Extracting Spatial Features

Once the neighborhood of each pixel has been adaptively defined, spatial information is extracted. Considering the small average size of the neighbors set, a description using higher order statistics would not be reliable. Shape descriptors are not appropriate either, as one given structure
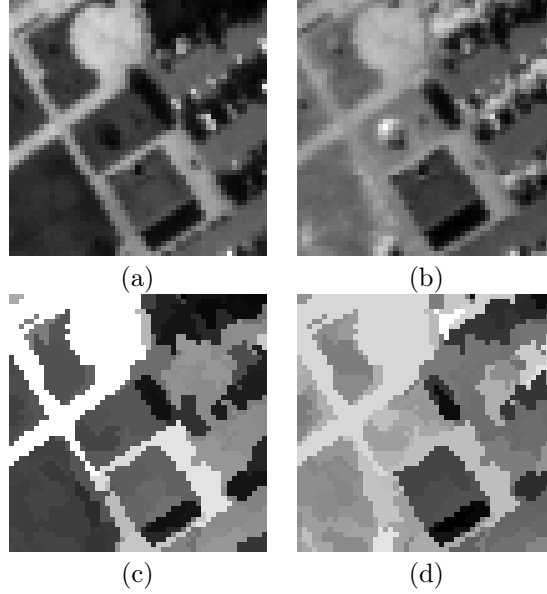
Figure 3: Original data and median filtering output ($\lambda = 15$): (a) band 50, (b) band 100, (c) and (d) filtered version of (a) and (b).

might be split into several consistent regions (see Figure 4.(b): The roof is divided into several triangles). As an alternative, the vector median value of the neighbors set $\Omega_{\mathbf{x}}$ is computed, for every pixel $\mathbf{x}$ [45]:

$$\Upsilon_{\mathbf{x}} = \mathrm{med}(\Omega_{\mathbf{x}}) \tag{1}$$

where $\dim(\mathbf{x}) = \dim(\Upsilon_{\mathbf{x}}) = n$, the number of spectral bands. Unlike the mean vector, the median vector is a vector from the initial set, which ensures a certain spectral consistency.

In order to speed up the kernel computation, the median filter can be applied in the pre-processing step, just after the area filtering. Then the inputs to the SVM classifier are the original data and the area+median filtered data. Figure 3 shows the results of such pre-processing on hyperspectral data.

In conclusion, every pixel now has two features: The *spectral feature* $\mathbf{x}$, which is the original value of each pixel, and the *spatial feature* $\Upsilon_{\mathbf{x}}$, which is the median value computed on each pixel's adaptive neighborhood. The easiest way to use both pieces of information would be to build a stacked vector, but it would not allow to weight the different features. In this paper, the kernel trick [13] of the SVM is exploited to design a composite kernel that makes it possible to set the relative influence of the extracted features. This is detailed in the next section.

3. Contextual Spatial-Spectral Support Vectors Machines

This section is dedicated to the definition of a contextual spatial-spectral SVM-based classifier. The classical setting is used to learn the SVM in the dual formulation [13]. Multiclass *One vs All* approach is chosen in this article, which consists in training an ensemble of classifiers, each

discriminating one class from all the others. The final decision is taken according to the classifier that provides the greatest distance from the hyperplane. This choice was motivated by the possibility of analyzing the results for each class separately, to investigate whether the neighborhood system defined above is well suited or not.

To apply SVM, one has to define a kernel function[1] between samples. For n-valued pixels:

$$k : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}. \tag{2}$$

One classical effective kernel is the Gaussian radial basis kernel:

$$k_\sigma(\mathbf{x}, \mathbf{z}) = \exp\left(-\frac{\|\mathbf{x} - \mathbf{z}\|^2}{2\sigma^2}\right) \tag{3}$$

where the norm is the Euclidean-norm and $\sigma \in \mathbb{R}^+$ tunes the variance of the Gaussian kernel. A short comparison of kernels for remotely sensed image classification can be found in [46].

For the classification of remote sensing images, $\mathbf{x}$ represents a pixel-vector where each component contains specific spectral information provided by a particular channel [47]. In this way a pixel-based classifier is defined. As explained in the introduction, the drawback is that inter-pixel dependency is not used.

Thanks to kernel properties, it is possible to define kernels that use both spectral and spatial information without running into intractable computational problems. Rules for kernel construction can be found in [48]. The linearity property is used to construct the new kernel: *If k_1 and k_2 are kernels, and $\mu_1, \mu_2 \geq 0$, then $\mu_1 k_1 + \mu_2 k_2$ is a kernel.*

Using the previous property, the spatial-spectral kernel $\mathcal{K}$ is defined as:

$$\begin{aligned} \mathcal{K}_{\sigma,\mu} \; : \; \mathbb{R}^n \times \mathbb{R}^n \; &\to \; [0,1] \\ (\mathbf{x}, \mathbf{z}) \; &\mapsto \; (1-\mu)k_\sigma^{spat}(\mathbf{x}, \mathbf{z}) + \mu k_\sigma^{spect}(\mathbf{x}, \mathbf{z}) \\ &\quad 0 \leq \mu \leq 1, \; 0 < \sigma \end{aligned} \tag{4}$$

From our experiments [46], the spectral kernel is defined as in (3):

$$\begin{aligned} k_\sigma^{spect} \; : \; \mathbb{R}^n \times \mathbb{R}^n \; &\to \; [0,1] \\ (\mathbf{x}, \mathbf{z}) \; &\mapsto \; \exp\left(-\frac{\|\mathbf{x} - \mathbf{z}\|^2}{2\sigma^2}\right) \end{aligned} \tag{5}$$

The spatial kernel is defined as follows:

$$\begin{aligned} k_\sigma^{spat} \; : \; \mathbb{R}^n \times \mathbb{R}^n \; &\to \; [0,1] \\ (\mathbf{x}, \mathbf{z}) \; &\mapsto \; \exp\left(-\frac{\|\Upsilon_\mathbf{x} - \Upsilon_\mathbf{z}\|^2}{2\sigma^2}\right) \end{aligned} \tag{6}$$

where $\Upsilon_\mathbf{x}$ is the spatial information extracted in Section 2, see eq.(1). Kernel (6) can be constructed by the composition of two functions: $k_\sigma^{spat} = k_\sigma \circ \Upsilon$ (where $\circ$ is a composition operator). The parameter σ is the same for both the spatial and the spectral kernel, the reason being the range of the data: Since the spatial features are extracted using median filtering from the spectral features, they have the same range. It is therefore appropriate to use the same parameter value.

---

[1] For convenience, in the following we refer to kernel instead of kernel function.

The weight μ controls the relative proportion of spatial and spectral information in the final kernel. For instance, for the class "Grass", the spectral information should be more discriminative while it should be the spatial information for the class "Building". This parameter is set during the training process, as the parameter σ. Finally, the new decision rule is:

$$f(\mathbf{z}) \quad = \quad \sum_{i=1}^{\ell} y_i \alpha_i \mathcal{K}_{\sigma,\mu}(\mathbf{x}_i, \mathbf{z}) + b. \tag{7}$$

For pixels belonging to the same set, the spatial information is the same. It is expected to achieve more homogeneous labeled zones in the final classification. This point is assessed in the next section.

4. Experimental Results

4.1. Data Sets

Three real data sets detailed below were used in the experiments. These data sets were selected for the following reasons:

1. All are very high spatial resolution images: This encourages the use of spatial information throughout the classification process.
2. Along with the high spatial resolution, two are hyperspectral data, *i.e.* with rich spectral information, and one is panchromatic image, *i.e.* with low spectral information.
3. They correspond to different scenarios: One peri-urban area and two dense urban areas.

The differences in terms of spatial or spectral resolutions and in terms of spatial structures present in the image cover a wide range of real cases. This allows to assess the robustness of the proposed method and the validity of the derived conclusions.

4.1.1. Hyperspectral Image

Airborne data from the ROSIS-03 (Reflective Optics System Imaging Spectrometer) optical sensor are used for the first two experiments. The flight over the city of Pavia, Italy, was operated by the Deutschen Zentrum für Luft- und Raumfahrt (DLR, the German Aerospace Agency) within the context of the HySens project, managed and sponsored by the European Union. According to specifications, the ROSIS-03 sensor provides 115 bands with a spectral coverage ranging from 0.43 to 0.86μm. The spatial resolution is 1.3 m per pixel. The two data sets are:

1. University Area: The first test set is around the Engineering School at the University of Pavia. It is 610×340 pixels. Twelve channels have been removed due to noise. The remaining 103 spectral channels are processed. Nine classes of interest are considered: Tree, asphalt, bitumen, gravel, metal sheet, shadow, bricks, meadow, and soil.
2. Pavia Center: The second test set is the center of Pavia. The image was originally 1096×1096 pixels. A 381 pixel wide black band in the left-hand part of image was removed, resulting in a 'two part' image of 1096×715 pixels. Thirteen channels have been removed due to noise. The remaining 102 spectral channels are processed. Nine classes of interest are considered: Water, tree, meadow, brick, soil, asphalt, bitumen, tile, and shadow.

Available training and test sets for each data set are given in Tables 1 and 2. These are pixels selected from the data by an expert, corresponding to predefined species/classes. Pixels from the training set are excluded from the test set in each case and vice-versa.

Table 1: Information classes and training-test samples for the University Area data set.

Class		Samples	
No	Name	Train	Test
1	Asphalt	548	6641
2	Meadow	540	18649
3	Gravel	392	2099
4	Tree	524	3064
5	Metal Sheet	265	1345
6	Bare Soil	532	5029
7	Bitumen	375	1330
8	Brick	514	3682
9	Shadow	231	947
	Total	3921	42776

Table 2: Information classes and training-test samples for the Pavia Center data set.

Class		Samples	
No	Name	Train	Test
1	Water	824	65971
2	Tree	820	7598
3	Meadow	824	3090
4	Brick	808	2685
5	Bare Soil	820	6584
6	Asphalt	816	9248
7	Bitumen	808	7287
8	Tile	1260	42826
9	Shadow	476	2863
	Total	7456	148152

4.1.2. Panchromatic Image

A panchromatic image is used in the third experiment. It comes from simulated PLEIADES (satellite to be launched in 2011) images provided by CNES (the French space agency). The image consists in 886×780 pixels. It has been acquired over the city of Toulouse, France. The spatial resolution is 0.75 meter per pixel and only one spectral band is available. Four classes were considered in each case, namely: Building, street, open area and shadow. See Table 3 for a description of the classes of interest and of the training/test sets.

4.2. Experiments

Three parameters need to be tuned for the SVM: the penalty term C, the width of the Gaussian kernel σ and the weight μ in $\mathcal{K}$. From previous works, C did not have a strong influence on the classification results when set greater than 10. For all the experiments, it was set at 200. The other two parameters were set using five-fold cross validation, $\sigma^2 \in \{0.5, 1, 2, 4\}$ and $\mu \in \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9\}$. The SVM algorithm was implemented with a modified version of LIBSVM [49]. Each original data set was scaled between $[-1, 1]$ using a per band range-stretching algorithm. In all experiments, the Gaussian radial basis kernel was used as the original kernel.

The proposed approach is compared to one state of the art approach, namely the Extended Morphological Profile (EMP) [42]. The EMP is constructed using a few first principal components of the multivalued images and by applying granulometries with geodesic filters. It has the limitations described in Section 2. In all the experiments, both the EMP and raw data were classified using the original SVM.

The classification accuracy was assessed with the overall accuracy (OA) which is the number of accurately classified samples divided by the number of test samples, the average accuracy (AA)

Table 3: Information classes and training-test samples for the PLEIADES data set.

	Class	Samples	
No	Name	Train	Test
1	Road	780	2450
2	Shadow	798	2588
3	Building	845	2293
4	Open Area	1738	3886
	Total	4161	11217

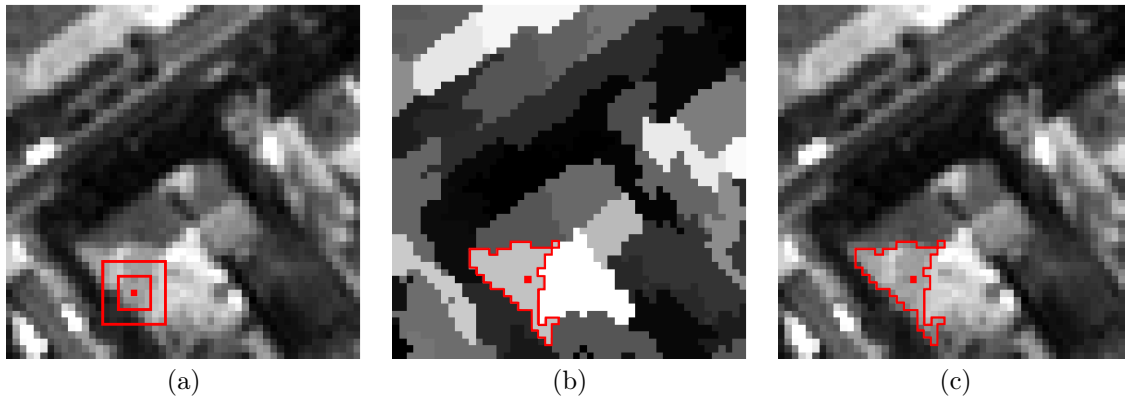|		
(a)	(b)	(c)

Figure 4: Inter-pixel dependency estimation. (a) Original image and fixed square neighborhood. (b) Filtered image and neighbor-set defined using area flat zones filter of size parameter $\lambda = 30$ [22]. (c) Original image with the defined neighbor-set $\Omega_{\mathbf{x}}$.

which represents the average of class classification accuracy, the kappa coefficient of agreement (κ) which is the percentage of agreement corrected by the amount of agreement that could be expected due to chance alone, and the class-specific accuracy. These criteria were used to compare classification results and were computed using the confusion matrix. Furthermore, the statistical significance of differences was computed using McNemar's test, which is based upon the standardized normal test statistic [50]:

$$Z = \frac{f_{12} - f_{21}}{\sqrt{f_{12} + f_{21}}} \tag{8}$$

where f_{12} indicates the number of samples classified correctly by classifier 1 and wrongly by classifier 2. The difference in accuracy between classifier 1 and 2 is said to be statistically significant if $|Z| > 1.96$. The sign of Z indicates whether classifier 1 is more accurate than classifier 2 ($Z > 0$) or vice-versa ($Z < 0$).

4.2.1. ROSIS University Area

Comparison with the original SVM. For this experiment, we first investigated the influence of the parameter λ on the definition of the neighborhood set. Several values for λ were tried, ranging from 2 to 40. Results of the classification are given in Table 4. Regarding the variation in the classification results, it seems that the parameter λ has a variable influence on the classification accuracies for each class. To explain this, three situations can be identified:

1. The spectral information is sufficient to discriminate the pixel, and this spectral information is not noisy, so no additional information is needed (classes 9 and 5).

12

Table 4: Classification accuracies for University Area data set. The best results for each class are reported in bold face. Δ is the difference between the best $\mathcal{K}^\lambda$ and the original kernel. $\mathcal{K}^\lambda$ means that classification was performed using the proposed kernel and area filtering of size λ.

Class	SVM	EMP	$\mathcal{K}^2$	$\mathcal{K}^5$	$\mathcal{K}^{10}$	$\mathcal{K}^{15}$	$\mathcal{K}^{20}$	$\mathcal{K}^{25}$	$\mathcal{K}^{30}$	$\mathcal{K}^{35}$	$\mathcal{K}^{40}$	Δ
1. Asphalt	80.64	**93.33**	80.41	80.08	82.39	83.25	84.57	86.32	84.36	84.57	86.13	5.68
2. Meadow	68.47	73.40	72.28	72.27	75.17	72.28	76.38	76.32	**78.52**	75.18	75.16	10.05
3. Gravel	73.80	52.45	76.51	77.13	81.71	87.04	**90.61**	89.71	84.80	85.52	85.90	16.81
4. Tree	97.49	**99.31**	96.44	96.64	94.19	94.94	95.07	94.61	96.87	94.84	97.65	0.16
5. Metal Sheet	99.49	99.48	99.11	99.26	99.41	96.51	99.48	98.29	**99.88**	99.63	99.63	0.39
6. Bare Soil	94.83	61.90	95.29	94.33	97.26	95.67	96.88	94.09	95.61	**98.31**	97.87	3.48
7. Bitumen	91.50	**97.67**	93.38	94.51	89.70	91.50	94.14	93.91	95.56	96.17	93.83	4.67
8. Brick	91.88	95.17	92.07	92.67	**96.14**	94.46	94.70	95.84	95.44	95.30	95.79	4.26
9. Shadow	97.04	92.29	95.04	94.83	95.56	92.29	93.56	93.24	**97.78**	96.73	97.25	0.74
OA	80.13	79.83	81.89	81.85	84.04	82.79	85.33	85.22	**86.11**	84.90	85.28	5.98
AA	88.33	85.00	88.95	89.09	90.17	89.77	91.71	91.37	91.98	91.80	**92.14**	3.81
κ	75.19	74.15	77.27	77.23	79.86	78.36	81.45	81.29	**82.35**	80.93	81.42	7.16

2. The size of the structure to which the sample belongs is:

   - large; when area filtering is performed, the sample is directly merged into a structure of a size larger than λ. This leads to a better discrimination of the concerned classes (classes 1, 7, and 8).

   - small; when area filtering is performed, the pixel may be merged into another structure. For example, when considering the class "Tree", this may be merged with class "Meadow". This leads to poorer discrimination of the concerned classes (class 4).

3. The class is highly textured and area filtering smoothes the structure. This leads to a better discrimination of the concerned classes (classes 2, 3, and 6).

According to the results in Table 4, the proposed approach outperforms the original SVM in terms of classification accuracies. The statistical significance of the differences in classification accuracy between the proposed approach and the original SVM are reported in Table 5. Kernel parameters found after the training step are given in Table 6. The value of μ confirms that a spatial kernel is useful for discrimination, since small values of μ are selected during the training process (corresponding to the inclusion of more spatial information than spectral information). It is worth noting that too high a value of λ does not help classification. It makes area filtering too strong, thereby removing too many relevant structures.

From this first experiment, it emerges that adding neighbors for the classification improved the final results for some classes but not for all. Regarding the nature of the classes, the optimum neighborhood system and the ratio between spatial and spectral information seem to be different for each class and need to be tuned during the training process.

Figure 5 shows a false color of the original image and the classification maps obtained using the original kernels and the proposed kernel.

Comparison with the Extended Morphological Profile [42]. Principal component analysis was applied to the data. The first three components were retained and morphological processing was applied. For each component, 4 opening/closings with a disk as structuring element, with initial radius of 2 and an increment of 2 were computed. Thus the EMP was a vector with 27 components. The parameters were fitted in the same manner as the classic kernel in the previous experiment.

Classification results for the EMP are listed in Table 4. The obtained difference statistic between our proposed approach and the EMP is $Z = 27.69$ with the best results obtained with $\lambda = 30$.

Table 5: Standardized Normal Test Statistic (Z) which Reflects the Significance of Differences in Classification Between Different Kernels for the University Area Data Set.

	SVM	$\mathcal{K}^{02}$	$\mathcal{K}^{05}$	$\mathcal{K}^{10}$	$\mathcal{K}^{15}$	$\mathcal{K}^{20}$	$\mathcal{K}^{25}$	$\mathcal{K}^{30}$	$\mathcal{K}^{35}$
$\mathcal{K}^{02}$	10.26								
$\mathcal{K}^{05}$	10.14	-0.38							
$\mathcal{K}^{10}$	22.66	15.51	15.86						
$\mathcal{K}^{15}$	15.29	5.17	5.43	-8.75					
$\mathcal{K}^{20}$	28.97	22.71	23.84	9.30	16.17				
$\mathcal{K}^{25}$	29.98	21.40	22.41	8.20	15.05	-1.06			
$\mathcal{K}^{30}$	33.30	26.84	27.66	14.37	20.11	5.85	7.35		
$\mathcal{K}^{35}$	27.33	18.55	18.89	5.98	13.6	-2.96	-2.26	-10.96	
$\mathcal{K}^{40}$	29.38	21.03	21.58	8.54	15.17	0.34	0.46	-7.12	3.78

Table 6: Kernel parameters found by 5-fold cross validation for University data set ($\lambda = 30$).

Class	1	2	3	4	5	6	7	8	9
μ	0.3	0.1	0.1	0.1	0.9	0.5	0.1	0.2	0.4
σ	0.5	1	0.5	0.5	1	1	1	0.5	4

From the table, the proposed approach performs better than the EMP in terms of classification accuracy. However, for the "Asphalt", "Tree" and "Bitumen" classes the EMP produces better classification. The "Asphalt" class corresponds to the roads in the image, which are typically thin, linear structures. Examining the thematic map Figure 5.(c), the roads seem to be better identified and more continuous than in Figure 5.(d). The morphological profile extracts information about the shape and size of a structure, while the median value of the adaptive neighborhood is more an indication of the gray-level distribution of a structure, hence it is not surprising that the MP classification performs best for that class. Note that both spatial-spectral approaches perform better for that class than the use of spectral information alone. For the "Tree" class, the interpretation is difficult. Two effects are to be considered: First, isolated trees are removed with the area filter thus making the proposed approach less efficient. Second, grouped trees are better classified with the proposed approach due to the smoothing effect. One effect can be more dominant than the other one, depending on the image.

In terms of classes with no typical shape such as meadow, gravel, or bare soil, the proposed approach outperforms the EMP in terms of classification accuracy. The adaptive neighborhood fits such "structures" better, and the value extracted from these "structures" helps in the discrimination (due to the smoothing effect described above).

In the next experiment, an image of a dense urban area is classified. According to the previous considerations mentioned above, the EMP ought to be the best in terms of classification accuracies at dealing with this type of data.

4.2.2. ROSIS Pavia Center

Comparison with the original SVM. For the second data set, classification accuracies were already very good. For convenience, only the best results have been reported (now λ is selected by cross-validation with the same range of variation as in the first experiment), compared with the original results, see Table 7. Following analysis of the parameter λ, it can be seen that classes 1, 2, 8, and 9 are separable using only the spectral information and adding spatial information does not help significantly in discrimination. Structures belonging to class 4 are merged together during the area filtering, and this leads to better discrimination. Textured classes (3, 5, and 7) are also better

<div align="center">(a)   (b)   (c)   (d)</div>

Figure 5: (a) False color original image of University Area. (b) Classification map using the RBF kernel. (c) Classification map using the EMP. (d) Classification map using the proposed kernel where $\lambda = 30$.

Table 7: Classification accuracies for Pavia Center data set for the standard SVM, the EMP and the proposed approach.

	OA	AA	κ	1. Water	2. Tree	3. Meadow	4. Brick	5. Bare Soil	6. Asphalt	7. Bitumen	8. Tile	9. Shadow
SVM	98.06	95.76	97.25	99.08	90.81	97.44	87.49	94.56	96.43	96.54	99.48	100
$\mathcal{K}^{20}$	98.43	97.13	97.79	99.15	90.04	98.12	94.00	99.45	95.82	98.15	99.47	99.93
EMP	98.95	97.72	98.51	99.82	90.94	95.50	99.07	99.06	97.99	97.49	99.62	99.97

separated. However, in this image, class 6 corresponds to narrow streets, and the area filtering impairs classification of this type of structure. Nevertheless, in the final analysis OA, AA and κ are improved with the proposed kernel, and the results are statistically different, with $Z = 8.82$.

Figure 6 presents the false color original image and classification maps using the original kernels and the proposed kernel.

Comparison with the EMP. The EMP was constructed following the same scheme as before. It comprised 27 features, and classification was performed using an SVM with a Gaussian kernel. Classification accuracies are reported in Table 7. The statistical difference is $Z = -15.81$, which means that the EMP performs better in this case. In terms of the global accuracy from the table, the EMP gave a slightly better result.

Regarding the class-specific accuracies, the same conclusions can be drawn as in the previous experiments. The class "Asphalt" is better classified using the EMP than with the proposed approach, while the class "Meadow" is better discriminated using the adaptive neighborhood. This confirms our previous conclusion. The class "Tile" was already accurately classified using spectral information, hence no significant difference is found between EMP and the proposed approach, though this class corresponds to most of the roofs in the image.

Figure 6 shows the thematic map obtained using the conventional SVM, the EMP and the proposed approach.
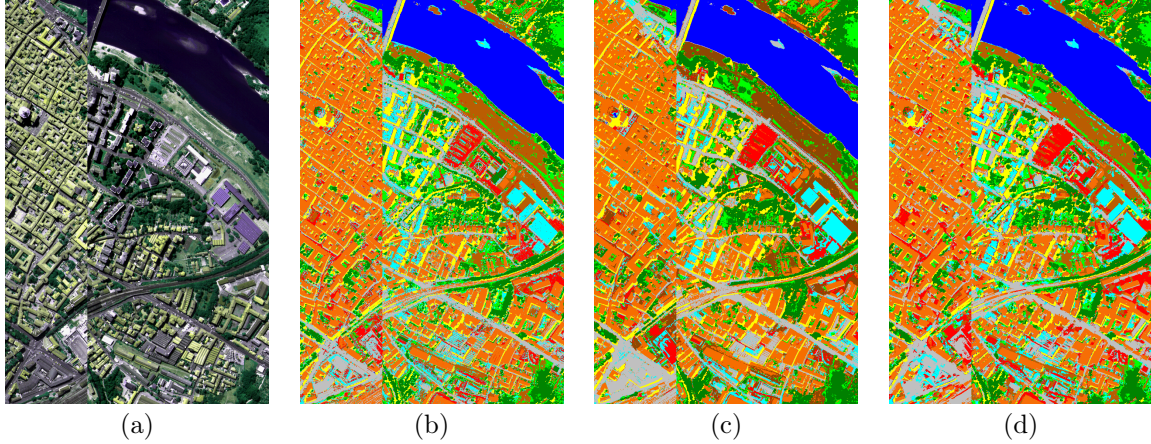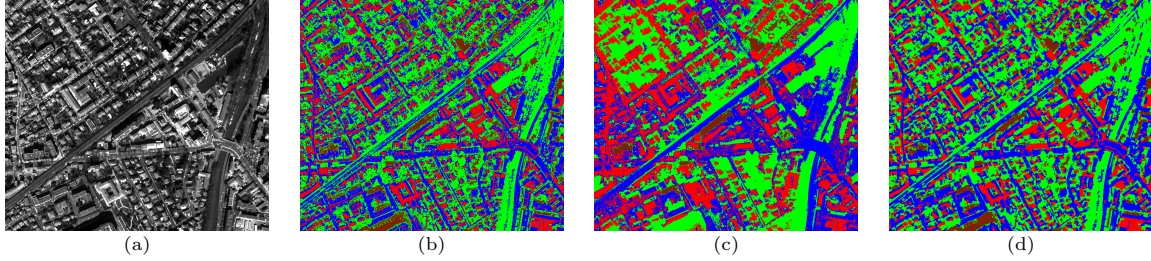
<div align="center">15</div>

Figure 6: (a) False color original image of Pavia Center. (b) Classification map using the RBF kernel. (c) Classification map using the EMP. (d) Classification map using the proposed kernel where $\lambda = 20$.

Table 8: Parameter μ found by cross-validation for PLEIADES Image

Class	1	2	3	4
μ	0.3	0.6	0.1	0.4

4.2.3. PLEIADES Image

Classification accuracies for the PLEIADES Image are reported in Table 9, as for the previous experiment 4.2.2, only the best results have been reported. The range of λ value was $[2, 60]$. The best CV result was obtained with $\lambda = 45$. The original image and the classification map using the original and the proposed kernel are shown in Figure 7. Classification accuracies have been significantly improved for each class.

Comparison with the original SVM. Classification accuracies are reported in Table 9. In terms of the three global classification accuracy estimates (OA, AA, and κ), the proposed approach led to an improvement in classification: The simultaneous use of spatial and spectral information invariably leads to better discrimination of the different classes.

In Table 8, the values of μ for each binary sub-problem are reported. The spatial information is given a heavier weight than the spectral information. This tends to prove that the proportion of each kind of information needs to be carefully tuned during the training process and should not be set to the same value for all classes.

Figure 7 shows the original image and the thematic maps obtained using the original kernel and the proposed kernel.

Comparison with the MP. Since the data are from a panchromatic image, the MP was used instead of the EMP (only one spectral band). It comprised 31 features, 15 geodesic openings/closings, the structuring element being a disk of size 2, 4...30. The MP leads to an improvement of the classification in terms of accuracy, but to a lesser extent than with the spatial-spectral SVM. From Figure 7, the thematic map produced with the MP features is less detailed than the thematic map produced with the proposed method.

16

Table 9: Classification accuracies for PLEIADES data set. $Z = 18.40$

	OA	AA	κ	1. Road	2. Building	3. Shadow	4. Open Area
SVM	76.99	75.53	68.33	70.69	86.09	60.97	84.35
$\mathcal{K}^{45}$	82.25	80.91	75.80	77.31	93.86	64.15	88.32
MP	80.58	78.55	73.14	82.42	48.03	85.52	98.15

(a)   (b)   (c)   (d)

Figure 7: (a) False color original image of Toulouse, (b) Classification map using the RBF kernel, (c) Classification map using the MP, (d) Classification map using the proposed kernel where $\lambda = 45$.

5. Discussion

5.1. General comments

Area Filtering. Used as a pre-processing step, this filter provides a simplified image where relevant structures are still present and details are removed. However, too high a value for λ may eliminate small structures, such as trees in the University Area data experiment or narrow streets in the Pavia Center data experiment. On the contrary, a too small value could make the method inefficient since in that case the defined neighborhood is too small. Hence this parameter needs to be chosen carefully. However, as it can be seen in Table 4, there is a range of value for which the algorithm performs equally well.

In the proposed approach, the parameter λ was selected globally, *i.e.*, all the classes share the same value. But in terms of classification accuracies, see Table 4, it can be seen that the optimum value of λ is class-dependent. Hence in future work, the parameter λ ought to be selected for each class independently.

Multiband Extension. Since there is no ordering relation between vector-valued pixels, direct extension of the area filter is not possible. The proposed approach is extended to hyperspectral data by considering the first principal component. This methodology has been successfully applied when using morphological filtering for the purpose of classification. In our experiments, the results confirm the interest of using this scheme.

Spatial-Spectral Kernel. This formulation allows a compact definition of the classification algorithm. Thus a few parameters (two) need to be tuned during the training stage. From these experiments, the values of σ does not seem to have a strong influence on the overall results ($\approx$ 1-2%) when the data are scaled between $[-1, 1]$. The relative proportion of spatial and spectral information has a stronger influence in the final classification.

A multi-resolution kernel can be defined using more than one spatial kernels: $\mathcal{K}_{\sigma,\mu} = \mu_0 k^{spect} + \sum_{i=1}^{d} \mu_i k_i^{spat}$ with $\sum_{i=0}^{d} \mu_i = 1$ and $\mu_i \geq 0$. For instance, the neighborhood can be defined at

different scales. Recent optimization algorithms, such as SimpleMKL [51], can be used to select the parameters μ automatically.

Spatial-Spectral SVM vs. EMP + SVM. In experiments, both approaches lead to improved classification in terms of accuracy. The spatial-spectral approach performs better for peri-urban areas, while the EMP leads to better results for very dense urban areas. However, the results are highly correlated to the definition of the classes: When considering classes according to geometrical characteristics (size or shape) the EMP performs better in terms of classification accuracies, but when considering classes according to textural or spectral characteristics, the spatial-spectral approach leads to better classification. Hence the method needs to be chosen in accordance with the data and the classes of interest. Note that area filtering is less demanding in terms of computing time than geodesic operators, especially for large images.

5.2. Computational complexity

Regarding the computational complexity, two main issues need to be considered, *i.e.*, the area filtering and the SVM classification. The area filtering represents at maximum 5% of the total processing time and is consequently negligible. The longer part is the parameter selection for the SVM (the cross validation). The computation of kernel (4) is about twice more demanding than the conventional Gaussian kernel. Still, it remains insignificant in comparison to running the SVM optimization problem. The complexity of the SVM is $o(n^3)$ where n is the number of training samples, which is the same whatever the method. Hence, once the kernel has been computed, the conventional SVM and the proposed method have almost the same complexity (the proposed method is slightly more demanding due to the kernel computation). However, there is one more parameter that needs to be selected (μ) for the proposed method, leading to a grid search by cross validation of 9×4 couples of values against only 4 for the conventional kernel (using the experimental settings). Thus, even if the complexity of the algorithms is almost the same, the parameters selection makes the proposed method take approximately 9 times longer than the conventional SVM method.

5.3. Comparison with other approach

Comparison with three others contextual methods for ROSIS-03 hyperspectral images are provided in this paragraph. The first method was proposed by S. Aksoy in [52]. He proposed to extract relevant spatial/spectral features using both spectral and spatial image filters (PCA, LDA and Gabor). Then a first supervised classification was performed using the extracted features and clusters are constructed based on the classifier output and a spatial morphological segmentation. The resulting features are characterized by spatial and spectral features and a supervised classification was performed on the cluster.

The second method was proposed by Y. Tarabalka *et al.* in [53]. It is based on a SVM classification followed by a Markov Random Field (MRF) regularization.

The third method was proposed by Huang and Zhang consist of a spatial multi-scale decomposition of the image followed by a classification with SVM [54]. The multi-scale decomposition was based on the mean-shift segmentation algorithm [55].

Overall accuracies are reported in Table 10. All the four methods perform very well for ROSIS-03 data sets. For the Pavia Center data set, the proposed method shows the best results in terms of classification accuracy together with the multi-scale approach. For the University Area, it is the SVM-MRF that performs the best and the second best result is provided by the proposed method. These comparisons confirm empirically the ability of the proposed method to classify hyperspectral images accurately.

18

Table 10: Classification accuracy of various methods on the ROSIS-03 data. Best results for each data sets are reported in boldface.

OA	Region level Bayesian [52]	SVM-MRF [53]	MS-SVM [54]	Proposed method
Pavia Center	96,7	97,6	**98,4**	**98,4**
University Area	84,5	**87,6**	82,7	86,1

6. Conclusion

The classification of remotely sensed images from urban areas using both spectral and spatial information has been considered. A key point is the definition of the spatial neighborhood and spatial information. In this article, the median value was used as a characterization of the inter-pixel dependency within a structure. Experiments have yielded good results in terms of classification accuracy. Defining a weighted kernel allows it to be applied with low computational complexity.

Comparisons were made between the proposed approach, the EMP and the original SVM. The proposed approach appears to perform better when the image area is not a dense urban area (University Area), while the EMP performs better for dense urban areas (Pavia Center). This is due to the morphological processing, which extracts geometrical information about the structure, while the proposed approach extracts only information about its gray-level distribution. However, in all cases, the proposed method outperforms the original SVM classifier.

One extension lies in the definition of the spatial information used. The median value does not provide information about the shape, size, or homogeneity of the neighborhood set. Other parameters could be extracted such as textural information, thus leading to another definition of the spatial kernel. In [56], a method for the estimation of the characteristic scale at each pixel was proposed. Such type of information needs also to be included in the classification process.

In future research, a connected filter should be defined and its influence on the definition of neighboring pixels has to be investigated. Multichannel area filters could also be addressed [57].

7. Acknowledgment

References

[1] S. Aksoy, N. H. Younan, L. Bruzzone, Pattern Recognition in Remote Sensing, Pattern Recognition Letters 31 (10) (2010) 1069 – 1070, ISSN 0167-8655, URL http://www.sciencedirect.com/science/article/B6V15-4YXK48T-2/2/58793f813d7d5520f0264bf1f1f7f84f.

[2] C. H. Chen, P.-G. Peter Ho, Statistical pattern recognition in remote sensing, Pattern Recognition 41 (9) (2008) 2731–2741, ISSN 0031-3203.

[3] D. L. Donoho, High-dimensional data analysis: the curses and blessing of dimensionality, in: AMS Mathematical challenges of the 21st century, 2000.

19

[4] K. Fukunaga, Introduction to statistical pattern recognition, CA: Academic Press, San Diego, 1990.

[5] G. F. Hughes, On the mean accuracy of statistical pattern recognizers, IEEE Trans. on Information Theory IT-14 (1968) 55–63.

[6] L. Jimenez, D. A. Landgrebe, Supervised classification in high dimensional space: geometrical, statistical and asymptotical properties of multivariate data, IEEE Trans. Syst., Man, Cybern. B 28 (1) (1993) 39–54.

[7] P. Soille, M. Pesaresi, Advances in Mathematical Morphology Applied to Geoscience and remote Sensing, IEEE Trans. Geosci. Remote Sens. 40 (9) (2002) 2042–2055.

[8] A. Plaza, J. A. Benediktsson, J. Boardman, J. Brazile, L. Bruzzone, G. Camps-Valls, J. Chanussot, M. Fauvel, P. Gamba, A. Gualtieri, J. C. Tilton, G. Triani, Recent Advances in Techniques for Hyperspectral Image Processing, Remote Sensing Environment (2009) S110–S122.

[9] D. A. Landgrebe, Signal Theory Methods in Multispectral Remote Sensing, John Wiley and Sons, New Jersey, 2003.

[10] X. Descombes, M. Sigelle, F. Preteux, GMRF Parameter Estimation in a non-stationary Framework by a Renormalization Technique: Application to Remote Sensing Imaging, IEEE Trans. Image Process. 8 (4) (1999) 490–503.

[11] Q. Jackson, D. A. Landgrebe, Adaptive Bayesian contextual classification based on Markov random fields, IEEE Trans. Geosci. Remote Sens. 40 (11) (2002) 2454–2463.

[12] J. A. Benediktsson, M. Pesaresi, K. Arnason, Classification and feature extraction for remote sensing images from urban areas based on morphological transformations, IEEE Trans. Geosci. Remote Sens. 41 (9) (2003) 1940–1949.

[13] V. Vapnik, Statistical Learning Theory, Wiley, New York, 1998.

[14] P. K. Varshney, M. K. Arora, Advanced Image Processing Techniques for Remotely Sensed Hyperspectral Data, Springer Verlag, 2004.

[15] B. Guo, S. Gunn, R. Damper, J. Nelson, Customizing Kernel Functions for SVM-Based Hyperspectral Image Classification, IEEE Trans. Image Process. 17 (4) (2008) 622–629.

[16] G. Camps-Valls, L. Gomez-Chova, J. Munoz-Mari, J. Vila-Francés, J. Calpe-Maravilla, Composite Kernels for Hyperspectral Image Classification, IEEE Geosci. Remote Sens. Lett. 3 (1) (2006) 93–97.

[17] G. Mercier, F. Girard-Ardhuin, Partially supervised oil-slick detection by SAR imagery using kernel expansion, IEEE Trans. Geosci. Remote Sens. 44 (10) (2006) 2839–2846.

[18] F. Bovolo, L. Bruzzone, M. Marconcini, A novel context-sensitive SVM for classification of remote sensing images, in: Proc IEEE Geoscience and Remote Sensing Symposium, IGARSS '06. Proceedings, 2498–2501, 2006.

[19] Z. Harchaoui, F. Bach, Image classification with segmentation graph kernels, in: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 1–8, 2007.

[20] S. Le Hegarat-Mascle, A. Kallel, X. Descombes, Ant colony optimization for image regularization based on a non-stationary Markov modeling, IEEE Trans. Image Process. 16 (3) (2007) 865–878.

[21] Y. Tarabalka, J. Chanussot, J. A. Benediktsson, Segmentation and classification of hyperspectral images using watershed transformation, Pattern Recognition 43 (7) (2010) 2367–2379, ISSN 0031-3203.

[22] P. Soille, Beyond self-duality in morphological image analysis, Image and Vision Computing 23 (2) (2005) 249–257.

[23] P. Soille, Morphological Image Analysis: Principles and Applications, Second Edition, Springer, Berlin, 2003.

[24] L. Bandeira, P. Pina, J. Saraiva, A multi-layer approach for the analysis of neighbourhood relations of polygons in remotely acquired images, Pattern Recognition Letters 31 (10) (2010) 1175 – 1183, ISSN 0167-8655, URL http://www.sciencedirect.com/science/article/B6V15-4YP0MRS-1/2/9cb785b79dd1ecb431d3966fa502dc24.

[25] Special Issue on Mathematical Morphology and Nonlinear Image Processing, Pattern Recognition 33 (6) (2000) 875 – 876, ISSN 0031-3203.

[26] J. Crespo, J. Serra, R. Schafer, Theoretical aspects of morphological filters by reconstruction, Signal Processing 47 (2) (1995) 201–225.

[27] J. Chanussot, J. A. Benediktsson, M. Fauvel, Classification of Remote Sensing Images from Urban Areas Using a Fuzzy Possibilistic Model, IEEE Geosci. Remote Sens. Lett. 3 (1) (2006) 40–44.

[28] M. Pesaresi, J. A. Benediktsson, A New Approach for the Morphological Segmentation of High-Resolution Satellite Imagery, IEEE Trans. Geosci. Remote Sens. 39 (2) (2001) 309–320.

[29] J. Chanussot, J. A. Benediktsson, M. Pesaresi, On the use of morphological alternated sequential filters for the classification of remote sensing images from urban areas, in: Proc. IEEE Geoscience and Remote Sensing Symposium, IGARSS '03. Proceedings, 473–475, 2003.

[30] M. Fauvel, J. Chanussot, J. Benediktsson, Adaptive pixel neighborhood definition for the classification of hyperspectral images with support vector machines and composite kernel, in: IEEE Conference on Image Processing (ICIP), 1884–1887, 2008.

[31] J. Crespo, R. W. Schafer, J. Serra, C. Gratin, F. Meyer, The flat zone approach: A general low-level region merging segmentation method, Signal Processing 62 (1) (1997) 37–60.

[32] P. Salembier, J. Serra, Flat zones filtering, connected operators, and filters by reconstruction, IEEE Trans. Image Process. 4 (8) (1995) 1153–1160.

[33] P. Monasse, F. Guichard, Fast Computation of a contrast-invariant image representation, IEEE Trans. Image Process. 9 (5) (2000) 860–872.

[34] L. Ikonen, Priority pixel queue algorithm for geodesic distance transforms, Image and Vision Computing 25 (10) (2007) 1520 – 1529.

[35] J. Debayle, J.-C. Pinoli, General Adaptive Neighborhood Image Processing - Part I, J. Math. Imaging Vis. 25 (2) (2006) 245–266.

[36] J. Debayle, J.-C. Pinoli, General Adaptive Neighborhood Image Processing - Part II, J. Math. Imaging Vis. 25 (2) (2006) 267–284, ISSN 0924-9907.

[37] R. M. Rangayyan, M. Ciuc, F. Faghih, Adaptive-Neighborhood Filtering of Images Corrupted by Signal-Dependent Noise, Appl. Opt. 37 (20) (1998) 4477–4487.

[38] J. Grazzini, P. Soille, Edge-preserving smoothing using a similarity measure in adaptive geodesic neighbourhoods, Pattern Recognition 42 (10) (2009) 2306 – 2316.

[39] E. Aptoula, S. Lefèvre, A comparative study on multivariate mathematical morphology, Pattern Recognition 40 (11) (2007) 2914 – 2929, ISSN 0031-3203.

[40] J. Chanussot, P. Lambert, Total ordering based on space filling curves for multivalued morphology, in: 4th International Symposium on Mathematical Morphology and its Applications, 51–58, 1998.

[41] P. Lambert, J. Chanussot, Extending mathematical morphology to color image processing, in: CGIP'00 - 1st International Conference on Color in Graphics and Image, 158–163, 2000.

[42] J. A. Benediktsson, J. A. Palmason, J. Sveinsson, Classification of hyperspectral data from Urban areas based on extended morphological profiles, IEEE Trans. Geosci. Remote Sens. 43 (3) (2005) 480–491.

[43] S. Valero, J. Chanussot, J. Benediktsson, H. Talbot, B. Waske, Advanced directional mathematical morphology for the detection of the road network in very high resolution remote sensing images, Pattern Recognition Letters 31 (10) (2010) 1120 – 1127, ISSN 0167-8655, URL http://www.sciencedirect.com/science/article/B6V15-4Y1VS9S-2/2/eb5c433a3e8fcd13cb2903e8e76abfa5.

[44] A. Plaza, P. Martinez, R. Perez, J. Plaza, A new approach to mixed pixel classification of hyperspectral imagery based on extended morphological profiles, Pattern Recognition 37 (6) (2004) 1097 – 1116.

[45] J. Astola, P. Haavisto, Y. Neuvo, Vector Median Filters, Proceedings of the IEEE 78 (4) (1990) 678–689.

[46] M. Fauvel, J. Chanussot, J. A. Benediktsson, Evaluation of kernels for multiclass classification of hyperspectral remote sensing data, in: IEEE International Conference on Acoustics, Speech and Signal Processing, vol. 2, ICASSP'06. Proceedings, 2006.

[47] C. Chang, Hyperspectral imaging. Techniques for spectral detection and classification, Kluwer Academic, 2003.

[48] B. Schölkopf, A. J. Smola, Learning with Kernels. Support vector machines, regularization optimization and beyond., MIT Press, 2002.

[49] C.-C. Chang, C.-J. Lin, LIBSVM: a library for support vector machines, software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm, 2001.

[50] G. M. Foody, Thematic map comparison: Evaluating the statistical significance of differences in classification accuracy, Photogrammetric Engineering & Remote Sensing 70 (5) (2004) 627–633.

[51] A. Rakotomamonjy, F. R. Bach, S. Canu, Y. Grandvalet, SimpleMKL, Journal of Machines Learning Research 9 (2008) 2491–2521.

[52] S. Aksoy, Signal and Image Processing for Remote Sensing, chap. 22, Saptial techniques for image classification, Taylor & Francis, 491–513, 2006.

[53] Y. Tarabalka, M. Fauvel, J. Chanussot, J. Benediktsson, SVM- and MRF-Based Method for Accurate Classification of Hyperspectral Images, IEEE Geosci. Remote Sens. Lett. 7 (4) (2010) 736 –740.

[54] X. Huang, L. Zhang, A comparative study of spatial approaches for urban mapping using hyperspectral ROSIS images over Pavia City, northern Italy, International Journal of Remote Sensing 30 (12) (2009) 3205 – 3221.

[55] D. Comaniciu, P. Meer, Mean Shift: A Robust Approach toward Feature Space Analysis, IEEE Trans. Pattern Anal. Mach. Intell. 24 (5) (2002) 603–619.

[56] B. Luo, J.-F. Aujol, Y. Gousseau, H. Maître, Resolution independent characteristic scale dedicated to satellite images, IEEE Trans. Image Process. 16 (10) (2007) 2503–2417.

[57] D. Brunner, P. Soille, Iterative area filtering of multichannel images, Image and Vision Computing 25 (8) (2007) 1352–1364.