

Spectral and Spatial Classification of Hyperspectral Data Using SVMs and Morphological Profiles

Mathieu Fauvel^{*†} *Student Member, IEEE*, Jón Atli Benediktsson[†] *Fellow, IEEE*, Jocelyn Chanussot^{*} *Senior Member, IEEE* and Johannes R. Sveinsson[†] *Senior Member, IEEE*

^{*}GIPSA-LAB, Signal & Image Department, Grenoble Institute of Technology - INPG
BP 46 - 38402 St Martin d'Herès - FRANCE

[†]Department of Electrical and Computer Engineering, University of Iceland
Hjardarhagi 2-6, 107 Reykjavík, ICELAND

E-mail: {mathieu.fauvel, jocelyn.chanussot}@gipsa-lab.inpg.fr, {benedikt, sveinsson}@hi.is

Abstract—A method is proposed for the classification of urban hyperspectral data with high spatial resolution. The approach is an extension of previous approaches and uses both the spatial and spectral information for classification. One previous approach is based on using several principal components from the hyperspectral data and building several morphological profiles. These profiles can be used all together in one extended morphological profile. A shortcoming of that approach is that it was primarily designed for classification of urban structures and it does not fully utilize the spectral information in the data. Similarly, the commonly used pixel-wise classification of hyperspectral data is solely based on the spectral content and lacks information on the structure of the features in the image. The proposed method overcomes these problems and is based on the fusion of the morphological information and the original hyperspectral data, i.e., the two vectors of attributes are concatenated into one feature vector. After a reduction of the dimensionality the final classification is achieved using a Support Vector Machines classifier. The proposed approach is tested in experiments on ROSIS data from urban areas. Significant improvements are achieved in terms of accuracies when compared to results obtained for approaches based on the use of morphological profiles based on PCs only and conventional spectral classification. For instance, with one data set, the overall accuracy is increased from 79% to 83% without any feature reduction and to 87% with feature reduction. The proposed approach also shows excellent results with a limited training set.

Index Terms—Data fusion, hyperspectral data, support vector machines, feature extraction, extended morphological profile, high spatial resolution.

I. INTRODUCTION

In classification of remote sensing data from urban areas, the identification of relatively small objects, e.g., houses and narrow streets is important. Therefore, high spatial resolution of the imagery is necessary for accurate classification. The most commonly available remote sensing data of high spatial resolution from urban areas are single-band panchromatic data. However, using only one high-resolution panchromatic data channel is usually not sufficient for accurate classification of structural information. To overcome that problem, Pesaresi and Benediktsson [1] proposed the use of morphological transformations to build a Morphological Profile (MP). In [2] the method in [1] was extended for hyperspectral data with high spatial resolution. The approach in [2] is based on using several Principal Components (PCs) from the hyperspectral data. From

each of the PCs, a morphological profile is built. These profiles are used all together in one Extended Morphological Profile (EMP), which is then classified by a neural network. The method in [2] has been shown to perform well in terms of accuracies when compared to more conventional classification approaches. However, a shortcoming of the approach is that it is primarily designed for classification of urban structures and it does not fully utilize the spectral information in the multispectral or hyperspectral data.

However this type of data contains a lot of information about the spectral properties and the land cover of the data. A finer definition of the classes is possible and more classes can be considered. Based on the spectral signatures of the classes, many advanced pixel-based classifiers have been proposed including advanced statistical classifiers [3] and distribution free approaches such as neural networks and support vector machines [4]. The later one has shown remarkable abilities to deal with remote multispectral data, especially with hyperspectral data. However, if the spatial content of the image is not used the resulting thematic map sometimes looks *noisy* (salt and pepper classification noise). Approaches involving Markov Random Field (MRF) and Monte Carlo optimization have been proposed in [5], [6]. These approaches use the *contextual information*. The main shortcoming of such algorithms is the computing time, which can be high even for small data sets. Regarding the high dimensionality of recently acquired data, both in the spectral and in the spatial domain, computationally light algorithm are of interest. In this sense, the MP has been proposed as an alternative way to use spatial information [1], [7]. Relatively to the MRF-based classifiers, the MP and its extension to a multiband image, the EMP, have the possibility to use geometrical contextual information (shape, size, *etc*) and perform well on many kinds of data (panchromatic, multispectral and hyperspectral data). However, as stated above, a shortcoming of this approach is it does not fully utilize the spectral information in the data, and consequently several approaches based on the MP/EMP have been proposed to fully exploit the spatial and the spectral information [8]–[10].

Each data set has its own properties, defining its ability to deal with different natures of classes. Table I sums up the properties of spectral and morphological/spatial data. The first main consideration is the complementary characteristics of the

data. It has a consequence in the discrimination ability of such a feature, as will be seen in the experiments. The fusion of two types of information should clearly results in an increase of the classification in terms of global accuracy. The use of spectral information can be critical for classification of non-structured information in urban areas, e.g., vegetation and soil classes while the use of spatial information can be useful for classification of structured objects, e.g., road and building classes.

The second consideration is the possible redundancy of each features set, see [3] for the spectral features and [11] for the spatial features. Hence feature extraction (FE) algorithms could be of an interest.

To include both type of information an extension to the approach in [2] is proposed in this paper. The proposed method is based on the data fusion of the morphological information and the original data: First, an extended morphological profile is created based on the PCs from the hyperspectral data. Secondly, feature extraction is applied on the original hyperspectral data and the extended morphological profile. Finally, the extracted feature vectors from both the original data and the extended morphological profile are concatenated into one stacked vector and classified. The proposed approach is different from approaches in [12]–[14], where the authors had extracted spatial information and used composite kernel to include both type of information. Here, feature extraction algorithms are used to select informative feature from the spectral and spatial domain.

For the multisource classification, Support Vector Machines (SVM) are used rather than a Neural Network, which was used in our previous experiment with MP/EMP. The superiority of SVM, implementing *structural risk minimization*, over the neural classifiers, implementing *empirical risk minimization*, has been discussed in [4] (in Chapters 9.6 and 12) and in [15], [16]. SVM aim to discriminate two classes by fitting an optimal separating hyperplane to the training data within a multi-dimensional feature space, by using only the closest training samples. Thus, the approach only considers samples close to the class boundary and works well with small training set, even when high dimensional data sets are classified. SVM have already been applied for multisource classification in [17] where several output coding methods were investigated.

In this paper, the proposed approach has been compared to statistical classification methods and SVM classification. Experiments were conducted on two different high resolution remote sensing data sets from urban areas. The effectiveness of the proposed methodology with a limited training set has been also assessed.

The paper is organized as follows. Section II reviews the use of morphological transformations for processing of hyperspectral imagery in urban areas. In Section III, the considered supervised feature extraction approaches are introduced. Support Vectors Machines (SVM) are discussed in Section IV. The applied data fusion schemes are discussed in Section V. Experimental results obtained on two ROSIS data sets from urban areas are presented in Section VI. Finally, conclusions are drawn in Section VII.

TABLE I
SPECTRAL AND SPATIAL DATA PROPERTIES. '↗' INDICATES A GOOD PROPERTY, '~' INDICATES THAT THE PROPERTY MIGHT BE HARMFUL AND '↘' INDICATES A CRITICAL PROPERTY.

Spectral features		Morphological features	
↗	Fine physical description	↗	Geometrical information
↗	Directly accessible	~	Needs to be extracted
~	Redundancy	~	Redundancy
↘	No spatial information	↘	Reduced spectral information

II. EXTENDED MORPHOLOGICAL PROFILE

Mathematical Morphology is a theory aiming to analyze spatial relationship between pixels. For a remote sensing application, several morphological operators are available for extracting geometrical information. An overview of operators can be found in [18]. In the following sub-section, some basic notions of mathematical morphology are reviewed. Then, the concept of the *Morphological Profile* and its extension to multivalued data are detailed.

A. Mathematical Morphology

The two fundamental operators in mathematical morphology are *erosion* and *dilation* [19]. These operators are applied to an image with a set of known shape, called a structuring element (SE). To erode an image consists of finding where the SE fits the objects in the image. The dilation, which is dual to the erosion, shows where the SE hits the objects.

Opening and *closing* are combinations of erosion and dilation. These operators remove from an original image structures of size less than the SE. But they also modify structures which are still present in the image after the opening/closing. Thus, they can introduce *fake* objects in the image. To avoid this problem, *geodesic morphology* and *reconstruction* should be used [19]. Opening and closing by reconstructions are connected operators that satisfy the following assertion: If the structure of the image cannot contain the SE, then it is totally removed, else it is totally preserved. For a given SE, geodesic opening or geodesic closing allows to know the size or shape of some objects present in the image: The objects that are smaller than the SE are deleted while the other (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*.

Granulometries are typically used for the analysis of the size distribution of the structures in the images. Classical granulometry by opening is built by successive opening operations with an SE of an increasing size. By doing so, the image is progressively simplified. Using connected operators, like opening by reconstruction, no shape noise is introduced.

Morphological Profiles (MPs) are defined using the granulometry. An MP is composed of the *opening profile* (OP) and the *closing profile* (CP). The OP at the pixel \mathbf{x} of the image I is defined as an n -dimensional vector:

$$OP_i(\mathbf{x}) = \gamma_R^{(i)}(\mathbf{x}), \quad \forall i \in [0, n] \quad (1)$$

where $\gamma_R^{(i)}$ is the opening by reconstruction with an SE of a size i and n is the total number of openings. Also, the CP at

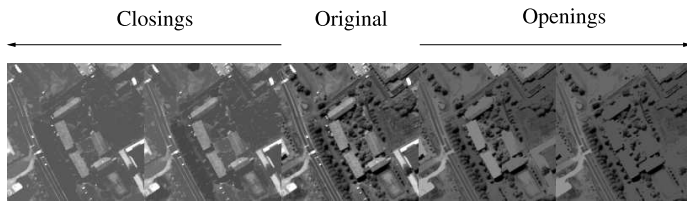


Fig. 1. Simple morphological profile with two opening and two closings. In the shown profile, circular structuring elements are used with radius increment 4 ($r = 4, 8$ pixels). The processed image is a part of Fig. 4(a).

the pixel \mathbf{x} of image I is defined as an n -dimensional vector:

$$CP_i(\mathbf{x}) = \phi_R^{(i)}(\mathbf{x}), \forall i \in [0, n] \quad (2)$$

where $\phi_R^{(i)}$ is the closing by reconstruction with an SE of a size i . Clearly we have $CP_0(\mathbf{x}) = OP_0(\mathbf{x}) = I(\mathbf{x})$. By collating the OP and the CP , the MP of image I is defined as $2n + 1$ -dimensional vector:

$$MP(\mathbf{x}) = \{CP_n(\mathbf{x}), \dots, I(\mathbf{x}), \dots, OP_n(\mathbf{x})\} \quad (3)$$

Example of MP is shown in Fig.1. Thus, from a single image results a multiband image, whose dimension corresponds to the number of transformations and spatial information is now contained in the MP for each pixel. However, an MP is built with only one band. Therefore, the spectral information is lost. One approach to deal with this problem is to extract several images that contain some parts of the spectral information and then build the MP on each of the individual images. This approach, namely the Extended Morphological Profile (EMP), is discussed in the following.

B. Extended Morphological Profile

In order to apply this approach to hyperspectral data, characteristic images need to be extracted. In [11], it was suggested to use several principal components of the hyperspectral data for such a purpose. Hence, the MP is applied on the first PCs, corresponding to a certain amount of the cumulative variance and a stacked vector is built with the MP on each PC. This yields to the *extended morphological profile* (EMP). Following the previous notation, the EMP is an $m(2n + 1)$ -dimensional vector:

$$MP_{ext}(\mathbf{x}) = \{MP_{PC^1}(\mathbf{x}), \dots, MP_{PC^m}\} \quad (4)$$

where m is the number of retaining PCs. Example of EMP is shown in Fig. 2.

As with multispectral data, the MP/EMP may include some redundancy. Classical feature reduction algorithm can be applied, as detailed in the following section.

III. SUPERVISED FEATURE EXTRACTION

Feature extraction can be viewed as finding a set of vectors that represents an observation while reducing the dimensionality. In pattern recognition, it is desirable to extract features that are focused on discriminating between classes of interest. Although a reduction in dimensionality is desirable, the error increment due to the reduction in dimension has to be without sacrificing the discriminative power of classifiers. In linear feature extraction, the number of input dimensions corresponds to the

number of selected eigenvectors [3]. The transformed data are determined by

$$\mathbf{y} = \Phi^T \mathbf{x} \quad (5)$$

where Φ is the transformation matrix composed of the eigenvectors of the feature matrix, \mathbf{x} is the data in the input space and \mathbf{y} is the transformed data in the feature space. We have in general $\dim(\mathbf{x}) \geq \dim(\mathbf{y})$. Several statistical extraction approaches have been proposed for remote sensing data [3], including *Decision Boundary Feature Extraction* (DBFE) and *Nonparametric Weighted Feature Extraction* (NWFE).

A. Decision Boundary Feature Extraction

It was shown in [20], that both discriminantly informative features and redundant features can be extracted from the decision boundary between two classes. The features are extracted from the decision boundary feature matrix (DBFM). The eigenvectors of the DBFM corresponding to non-zero eigenvalues are the necessary feature vectors to achieve the same classification accuracy as in the original space. The efficiency of the DBFE is related to the training set and can be computationally intensive.

B. Nonparametric Weighted Feature Extraction

To overcome the limitations of the DBFE, Kuo and Landgrebe [21] proposed the nonparametric weighted feature extraction. NWFE is based on the Discriminant Analysis Feature Extraction by focusing on samples near the eventual decision boundary. The main ideas of the NWFE are 1) putting different weights on every sample to compute the local means and 2) defining nonparametric between-class and within-class scatter matrices [3].

Many experiments have shown the effectiveness of these approaches for the classification of hyperspectral data [3]. They are usually applied on the spectral data, but it was successfully applied to the EMP [11].

IV. CLASSIFICATION BY THE SUPPORT VECTOR MACHINE

So far, in our previous approach [2], [7], [11], [22] the classification was done with either a statistical classifier (Gaussian Maximum Likelihood), a neural network or a fuzzy classifier. Here it is proposed to use the *Support Vector Machines* (SVM). Early work in classification of remotely sensed images by SVM showed excellent results [17], [23], [24]. In [15], several SVM-based classifiers were compared to other classical classifiers such as a K-nearest neighbors classifier and a neural network classifier and the SVM using the kernel method outperformed the other classifiers in terms of accuracy. Multiclass SVM performances were also positively compared with a discriminant analysis classifier, a decision tree classifier and a feedforward neural network classifier with a limited training set [25]. SVM show good results in the situation of limited training set in [26]. Semisupervised SVM were also investigated for multi-spectral data classification [27], [28].

SVM are surely among the most used kernel learning algorithm. It performs robust non-linear classification of samples using the kernel trick. The idea is to find a separating hyperplane

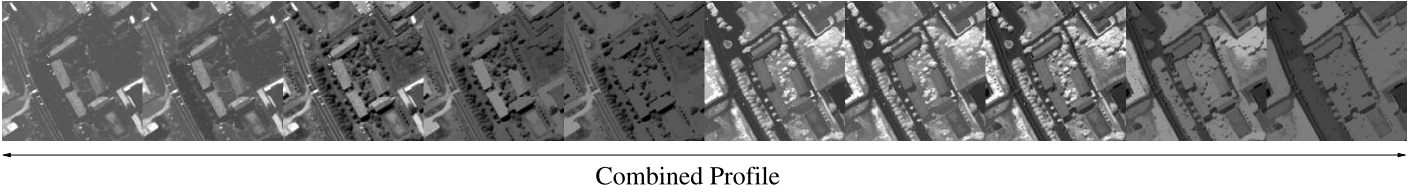


Fig. 2. Extended morphological profile of two images. Each of the original profile has two openings and two closings. Circular structuring element with radius increment 4 was used ($r = 4, 8$). The processed image is a part of Fig. 4.(a).

in some feature space induced by the kernel function while all the computations are done in the original space [4]. A good introduction to SVM for pattern recognition can be found in [29]. Given a training set $\mathcal{S} = \{(\mathbf{x}^1, y_1), \dots, (\mathbf{x}^\ell, y_\ell)\} \in \mathbb{R}^n \times \{-1; 1\}$, the decision function is found by solving the convex optimization problem:

$$\begin{aligned} \max_{\alpha} g(\alpha) &= \sum_{i=1}^{\ell} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{\ell} \alpha_i \alpha_j y_i y_j k(\mathbf{x}^i, \mathbf{x}^j) \\ \text{subject to} & \quad 0 \leq \alpha_i \leq C \text{ and } \sum_{i=1}^{\ell} \alpha_i y_i = 0 \end{aligned} \quad (6)$$

where α are the Lagrange coefficients, C a constant that is used to penalize the training errors, and k the kernel function. To be an acceptable kernel, k should be a positive semi-definite function [30]. One classical effective kernel is the Gaussian kernel:

$$k_{\sigma}(\mathbf{x}^i, \mathbf{x}^j) = \exp\left(-\frac{\|\mathbf{x}^i - \mathbf{x}^j\|^2}{2\sigma^2}\right) \quad (7)$$

where the norm is the *Euclidean*-norm and $\sigma \in \mathbb{R}^+$ tunes the flexibility of the kernel. A short comparison of kernels for remotely sensed image classification can be found in [26].

When the optimal solution of (6) is found, i.e., the α_i , the classification of a sample \mathbf{x} is achieved by looking to which side of the hyperplane it belongs:

$$y = \text{sgn}\left(\sum_{i=1}^{\ell} \alpha_i y_i k(\mathbf{x}^i, \mathbf{x}) + b\right). \quad (8)$$

To deal with multiclass classification problem, the pairwise approach was used in our experiments [31]. More advanced multiclass approaches applied to remote sensing data can be found in [15]. For the particular case of one-class-classification, a dedicated methodology is proposed in [32].

The SVM are mainly a *non-parametric* method, yet some parameters need to be tuned before the optimization. In the Gaussian kernel case, there are two parameters: C , the penalty term, and σ , the width of the exponential. It is usually done by a cross-validation step, where several values are tested. In our experiments, C was fixed to 200 and $\sigma^2 \in \{0.5, 1, 2, 4\}$ was selected using a 5-fold cross validation. The SVM optimization problem was solved using the LIBSVM [33].

V. DATA FUSION

The proposed method is based on the data fusion of the morphological information and the original data. In a previous work [34], it was proposed to fuse the classification results of two SVM classifiers, each one working with either the spectral or the EMP data. It consisted in an appropriate adaptive fusion

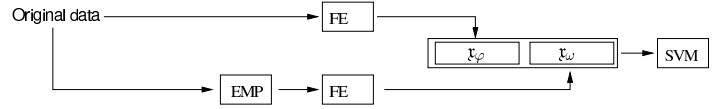


Fig. 3. Proposed data fusion scheme.

scheme based on the output's characteristics of the SVM. The results in terms of accuracy were increased but it needed two training of SVM, that could be time consuming.

Here it is proposed to use a multisource strategy to fuse spectral and spatial information. First, an EMP is created based on applying the PCA on the hyperspectral data. Secondly, feature extraction is applied on both the EMP and the original hyperspectral data. Finally, the extracted feature vectors are concatenated into one stacked vector and classified by the SVM.

In the morphological processing we usually retain PCs corresponding to 99% of the cumulative variance. This is done in order to reduce the redundancy in the data but keep most of the variation. The EMP is built using the m PCs that correspond to the 99% variance. Each MP is composed of n geodesic openings, n geodesic closing and the corresponding PC. The SE is a disk with initial radius of r pixels. The size increment is s . Hence, each MP has $2n + 1$ features and the EMP has $m(2n + 1)$ features. Noting \mathbf{x}_{ϕ} the features associated to the spectral bands and \mathbf{x}_{ω} the features associated to the EMP, the corresponding extracted features from the FE algorithm are:

$$\mathbf{r}_{\phi} = \Phi_{\phi}^T \mathbf{x}_{\phi} \quad (9)$$

and

$$\mathbf{r}_{\omega} = \Phi_{\omega}^T \mathbf{x}_{\omega}. \quad (10)$$

The stack vector is finally $\mathbf{r} = [\mathbf{r}_{\phi}, \mathbf{r}_{\omega}]^T$.

Fig. 3 presents the data fusion scheme. Note that in this work, only morphological information is extracted, but it is possible to extract other types of spatial information with other processing and include them in the stacked vector.

VI. EXPERIMENTS

A. Data set

Airborne data from the ROSIS-03 (Reflective Optics System Imaging Spectrometer) optical sensor are used for the experiments. The flight over the city of Pavia, Italy, was operated by the Deutschen Zentrum für Luft- und Raumfahrt (DLR, the German Aerospace Agency) in the framework of the HySens project, managed and sponsored by the European Union. According to specifications, the number of bands of the ROSIS-03 sensor is 115 with a spectral coverage ranging from 0.43 to

TABLE II
INFORMATION CLASSES AND TRAINING-TEST SAMPLES FOR THE
UNIVERSITY AREA DATA SET.

Class		Samples	
No	Name	Train	Test
1	Asphalt	548	6641
2	Meadows	540	18649
3	Gravel	392	2099
4	Trees	524	3064
5	Metal Sheets	265	1345
6	Soil	532	5029
7	Bitumen	375	1330
8	Bricks	514	3682
9	Shadows	231	947
Total		3921	42776

TABLE III
INFORMATION CLASSES AND TRAINING-TEST SAMPLES FOR THE PAVIA
CENTER DATA SET.

Class		Samples	
No	Name	Train	Test
1	Water	824	65971
2	Trees	820	7598
3	Meadows	824	3090
4	Bricks	808	2685
5	Soil	820	6584
6	Asphalt	816	9248
7	Bitumen	808	7287
8	Tiles	1260	42826
9	Shadows	476	2863
Total		7456	148152



(a)

(b)

Fig. 4. ROSIS data: (a) University Area, (b) Pavia Center. Three-channel color composite of the areas used for the classification. Data characteristics are detailed in Sub-Sections VI-B and VI-C.

age of agreement corrected by the amount of agreement that could be expected due to chance alone [35].

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 [36]:

$$Z = \frac{f_{12} - f_{21}}{\sqrt{f_{12} + f_{21}}} \quad (11)$$

where f_{12} indicates the number of samples classified correctly by classifier 1 and incorrectly by classifier 2. The difference in accuracy between classifiers 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 ($Z < 0$). This test assumes related testing samples and thus is adapted to our situation since the training and testing set were the same for each experiment.

The feature extractions were done with MultiSpec© [3] while the morphological operations were done with the Image Processing Toolbox of Matlab©. The SVM classification was done using the LIBSVM through its Matlab© interface [33]. From previous experiments on the same data set, the Gaussian kernel provides the best results and was used for the experiments [26]. The range of each feature, be it spectral or morphological, was stretched between 0 and 1.

To obtain a baseline result for comparison, the classification was also done using the Gaussian Maximum Likelihood (ML) classifier on the hyperspectral data using MultiSpec©¹. Feature extraction was done using the two FE algorithms (DBFE and NWFE) but only the best results have been reported for both data sets. The results were compared to those obtained by the proposed approach.

¹For the ML, the kappa coefficient was not accessible.

0.86 μ m. The data have been atmospherically corrected but not geometrically corrected. The spatial resolution is 1.3m per pixel. Two data sets were used in the experiment:

- 1) *University Area*: The first test set is around the Engineering School at the University of Pavia. It is 610 by 340 pixels. Some channels (twelve) have been removed due to noise. The remaining 103 spectral dimensions are processed. Nine classes of interest are considered, i.e., trees, asphalt, bitumen, gravel, metal sheets, shadow, bricks, meadows and soil.
- 2) *Pavia Center*: The second test set is the center of Pavia. The Pavia center image was originally 1096 by 1096 pixels. A 381 pixel wide black stripe in the left part of image was removed, resulting in a "two parts" image. This "two parts" image is 1096 by 715 pixels. Some channels (thirteen) have been removed due to noise. The remaining 102 spectral dimensions are processed. Nine classes of interest are considered, i.e., water, trees, meadows, bricks, soil, asphalt, bitumen, tiles and shadows.

Available training and testing set for each data set are given in Table II and III and Fig. 4 presents false colors images for both data set.

The classification accuracy was assessed with:

- An overall accuracy(OA) which is the number of well classified samples divided by the number of test's samples
- An average accuracy (AA) which represents the average of class classification accuracy
- A kappa coefficient of agreement (κ) which is the percent-

TABLE IV
UNIVERSITY AREA. EIGENVALUES OF PRINCIPAL COMPONENTS IN PERCENTAGE.

	Value	Cumulative Val.
λ_1	82.75	82.75
λ_2	15.56	98.01
λ_3	1.51	99.55
λ_4	0.14	99.70

B. University Area data set

PCs were computed from the hyperspectral data. The results for the eigenvalues are shown in Table IV. The left column gives the component number, the center column the eigenvalues in percentage of the total amount of variance and the right column the cumulative amount of variance. From the table, three PCs were necessary to retain 99% of the variance criterion. EMPs were built according to the scheme presented in the Section V: A circular SE with a step size increment of 2 was used. Four openings and closings were computed for each PC, resulting in an EMP of dimension 27.

First, the classification with SVM was done using the spectral information and the extended morphological profile. The best ML accuracy was obtained using 8 features extracted with the NWFE, following Landgrebe's recommendations in [3]. The results are reported in Table V. Regarding the global accuracies, both SVM approaches perform equally well, for instance the difference between the classification using the spectral information and the EMP is not statistically significant, see Table VI. Note that it is consistent with the characteristics of the scene: The University Area is a mix between man-made structures and natural materials. Therefore, the morphological information is not as useful as it could be in a very dense urban area. When a careful analysis is done on the class-specific accuracies, we can see from Table V that each approach performed well for complementary classes, e.g. the spectral approach performed better for classes 3, 6, 9 while the EMP approach performed better for classes 1, 2, 7, 8. After the data fusion we have to look at these classes and see if the best information was used, i.e., if the classification accuracy increased for these classes.

The experiment was then performed with the concatenated vector. The vector was made of the 103 spectral bands and the 27 features of the EMP. The vector was directly used as an input to the SVM. The classification results are reported in Table V. As can be seen from the table, the global accuracies increased. The κ value in percentage is 79.13% against 74.47% for the spectral approach and 73.25% for the EMP and the differences are statistically significant (see Table VI). Regarding the class-specific accuracies, the results in terms of accuracies have increased for classes 1, 7, 8 when compared to both individual approaches. In fact all the classes are more accurately classified than the worst respective cases for the individual approaches.

In the last experiment, feature reduction was applied on the morphological data and the original data before the concatenation. Then the stacked vector was classified by the SVM. Table VII summarizes the test accuracies for several values of the variance criterion for the DBFE and NWFE. Best results were obtained with 95% and 80% variance criterion for the

TABLE VII
UNIVERSITY AREA. GLOBAL ACCURACIES IN PERCENTAGE WITH DIFFERENT FEATURE EXTRACTION METHODS. THE NUMBERS OF FEATURES FROM THE SPECTRAL DATA AND THE MORPHOLOGICAL DATA, RESPECTIVELY, ARE GIVEN IN BRACKETS.

Feature extraction	Cum. Var.	Features	OA	AA	κ
DBFE	99	60 (45,15)	84.77	89.98	80.43
	95	27 (27,10)	87.97	88.94	84.40
	90	28 (20,8)	86.49	88.94	82.50
	80	19 (14,5)	82.95	87.51	77.27
	70	14 (10,4)	76.93	84.63	71.20
NWFE	99	62 (42,20)	84.15	88.89	79.61
	95	28 (16,12)	82.90	87.25	77.98
	90	18 (10,8)	82.64	86.77	77.65
	80	13 (7,6)	87.59	88.93	83.89
	70	10 (5,5)	79.57	87.39	74.49

DBFE and NWFE, respectively. Using 95% of the variance criterion with DBFE, the hyperspectral data were reduced to 27 features and the EMP to 10 features. With NWFE and 80%, 7 features were extracted from the hyperspectral data and 6 from the EMP. Again, as can be seen in Table VI, differences between the classification accuracies are statistically significant.

Considering the class-specific accuracies, the DBFE approach improved the classification for class 2 while class 3 was less accurately classified than with the concatenated full hyperspectral data and EMP. However, the DBFE outperformed the individual classifications of the spatial or spectral information. On this data set, the classification of the DBFE feature extracted data gave the best classification results. Similar comments can be made for the accuracies obtained with classifications of the NWFE. Still, the number of features needed to achieve the same accuracy is significantly lower for the NWFE approach than for the DBFE. Since the SVM is linearly related to the dimensionality of the data, lower dimensional data reduced the training time and increased the speed of the classification.

To assess this increase, comparison of the processing time (training and classification process) for the different approaches was made. Table VIII summarizes the results, which are clearly different according to the features used. The training time could depend on several factors:

- 1) The dimension of the data;
- 2) The size of the training set;
- 3) The number of parameters for the kernel.

For our given problem, items 2) and 3) are the same. Reducing the size of the data is beneficial for the processing time, since data with lower dimensionality (EMP and NWFE) have the shortest processing time. For the best case (NWFE), the gain is about 73%.

For the classification processing time, two factors have an influence: The dimension of the data and the number of support vectors (non-zero α_i in (8)). Thus, approaches with low dimensionality and few support vectors perform the classification task of the whole image faster (EMP and NWFE). Nevertheless, the classification processing is really fast by comparison to the training time, in all the cases.

Classification maps for the different approaches are shown in Fig. 5.

TABLE V
UNIVERSITY AREA. SUMMARY OF THE GLOBAL AND THE CLASS-SPECIFIC TEST ACCURACIES IN PERCENTAGE FOR THE CLASSIFICATION. THE NUMBERS OF FEATURES FROM THE SPECTRAL DATA AND THE MORPHOLOGICAL DATA, RESPECTIVELY, ARE GIVEN IN BRACKETS.

	ML (NWFE)	Spectral	EMP	Spec. EMP	DBFE 95%	NWFE 80%
Features	8	103	27	130	37 (27,10)	13 (7,6)
OA	80.10	79.48	79.14	83.53	87.97	87.59
AA	87.00	88.14	84.30	89.39	89.43	88.93
κ	-	74.47	73.25	79.13	84.40	83.89
Class 1	76.00	84.36	94.50	95.33	90.92	86.80
Class 2	73.90	66.20	72.82	73.46	85.91	86.95
Class 3	70.80	71.99	53.22	65.89	57.88	63.26
Class 4	96.70	98.01	98.89	99.18	99.22	98.53
Class 5	99.90	99.48	99.55	99.48	99.48	99.88
Class 6	87.60	93.12	58.11	84.15	85.32	82.62
Class 7	92.00	91.20	96.09	97.22	95.19	96.61
Class 8	87.30	92.26	95.27	96.12	95.84	95.38
Class 9	99.10	96.62	91.24	93.66	95.14	90.60

TABLE VI
UNIVERSITY AREA. STATISTICAL SIGNIFICANCE OF DIFFERENCES IN CLASSIFICATION ACCURACIES.

	EMP/Spectral	Spectral/Spec. EMP	Spectral/DBFE 95%	Spectral/NWFE 80%	Spec. EMP/DBFE 95%	Spec. EMP/NWFE 80%	DBFE 95%/NWFE 80%
Z	-1.36	-19.55	-37.98	-36.61	-25.19	-21.79	2.32

TABLE VIII
UNIVERSITY AREA. PROCESSING TIME IN SECONDS AS FUNCTION OF DIMENSIONALITY AND NUMBER OF SUPPORT VECTORS.

	Spectral	EMP	Spec. EMP	DBFE 95%	NWFE 80%
Dimension	103	27	130	37	13
Training	3074	850	3257	1184	859
Number of SVs	1085	406	529	727	572
Classification	76	9	41	18	9

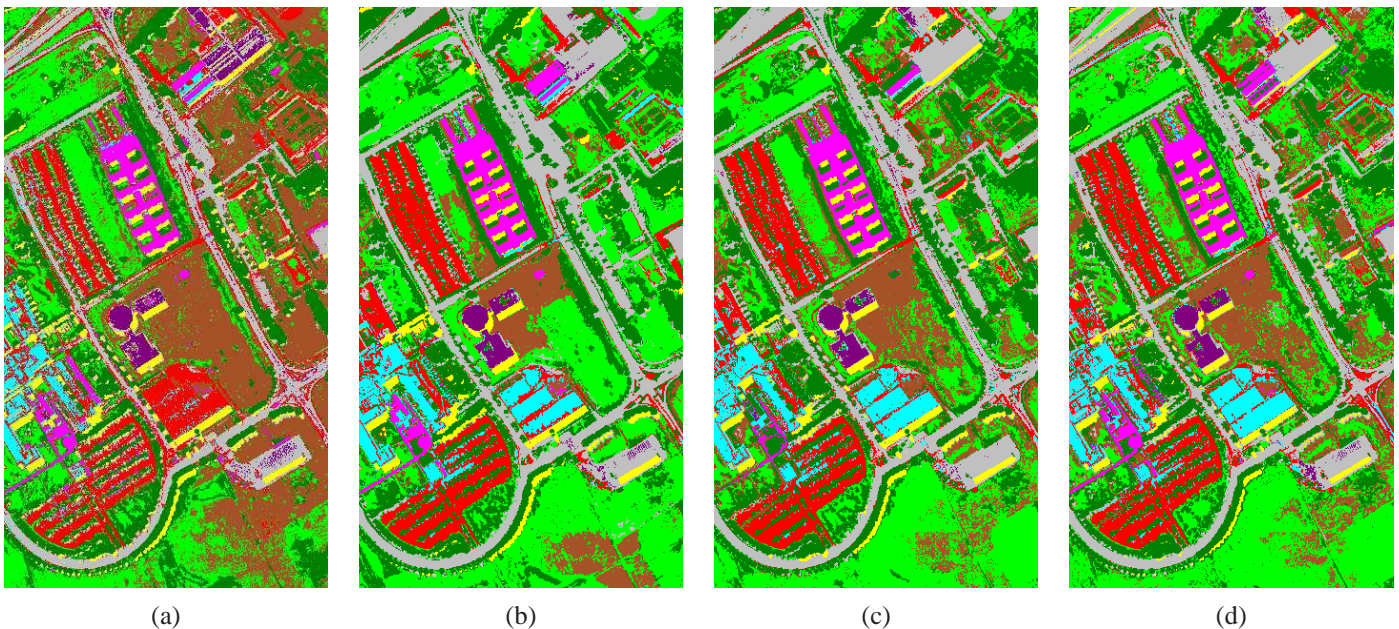


Fig. 5. University Area: Classification map obtained with SVM from: (a) the original hyperspectral data, (b) the EMP, (c) 37 DBFE features and (d) 13 NWFE features. Classification accuracies are reported in the Table V.

C. Pavia Center data set

For the second test, the scene is a very dense urban area in the center of the city of Pavia. Because of that, morphological information should be useful for the discrimination. PCs were computed from the hyperspectral data. The results for the

eigenvalues are shown in Table IX. From the Table, three PCs were necessary to retain 99% of the variance criterion. The EMP was built according to the scheme presented in the section V: A circular SE with a step size increment of 2 was used. Four openings and closings were computed for each PC, resulting in

TABLE IX
PAVIA CENTER. EIGENVALUES OF PRINCIPAL COMPONENTS IN PERCENTAGE.

	Value	Cumulative Val.
λ_1	82.94	82.94
λ_2	14.82	97.77
λ_3	1.70	99.47
λ_4	0.19	99.66

an EMP of dimension 27.

SVM classification was applied to the original hyperspectral data and the EMP. The best ML accuracy was obtained using 29 features extracted with the DBFE. The results are reported in Table X. From the Table, it can be seen that SVM classifier achieved excellent global accuracies. In these experiments, the morphological approach performs better than the spectral based approach. Table XI shows the statistical significance of differences between the classification accuracies for the different approaches. This is consistent with the characteristics of the picture: it is a very dense urban area and morphological processing provides discriminative information. In terms of accuracies, the main improvement in the classification is achieved for class 4. The other classes are classified equally accurately. The data fusion should thus improve the classification of class 4 while preserving very good results for the others classes.

Next, the experiment was performed using the concatenated vector. The vector was made of the 102 spectral bands and the 27 features of the EMP. This vector was used as an input for the SVM without any additional processing. The classification results are reported in Table X. The differences of classification accuracies between the EMP and the concatenated vector are not statistically significant, since the McNemar's test is almost equal to zero, see Table XI. Thus, both EMP and concatenated vector perform equally well.

As in the previous experiment, feature reduction was applied both on the morphological data and on the original data before the concatenation. Then, the stacked vector was classified by the SVM. Table XII summarizes the test accuracies for several values for the variance criterion for the DBFE and NWFE. The best results are obtained with 99% variance criterion for both DBFE and NWFE. Using 99% of the variance with the DBFE, the hyperspectral data is reduced to 51 features and the EMP is reduced to 15 features. With the NWFE and 99% of the variance criterion, 44 features were extracted from the hyperspectral data and 20 from the EMP. The results are given in Table X.

For this experiment, the DBFE does not help for the classification since the Z test is not significant. On the other hand, similar classification accuracy is reached with far less features, nearly half the size of the previous feature set, thus decreasing the total training and classification time. The NWFE leads to a significant increase of the classification accuracies, $|Z|=7.75$ by comparison to the best results obtained with the concatenation vector, which is contrary to the previous experiment. Classification maps for the different approaches are shown in Fig. 6. Visually, the thematic map produced with the classification of the NWFE features seems less noisy than the one obtained with the classification of the DBFE features. This is especially true in the top-left corner which correspond to a

TABLE XII
PAVIA CENTER. GLOBAL ACCURACIES IN PERCENTAGE WITH DIFFERENT FEATURE EXTRACTION METHODS. THE NUMBERS OF FEATURES FROM THE SPECTRAL DATA AND THE MORPHOLOGICAL DATA, RESPECTIVELY, ARE GIVEN IN BRACKETS.

Feature extraction	Cum. Var.	Features	OA	AA	κ
DBFE	99	66 (51,15)	98.65	97.30	98.10
	95	41 (31,10)	98.37	96.86	97.70
	90	31 (23,8)	98.08	96.71	97.29
	80	22 (17,5)	98.53	97.22	97.42
NWFE	70	17 (13,4)	98.53	97.31	97.42
	99	66 (44,20)	98.87	97.95	98.41
	95	31 (19,12)	98.58	96.66	97.99
	90	21 (12,9)	98.41	97.28	97.71
	80	14 (8,6)	98.24	96.63	97.52
70	10 (6,4)	98.39	96.39	97.73	

TABLE XIV
UNIVERSITY AREA. SUMMARY OF THE GLOBAL TEST ACCURACIES IN PERCENTAGE FOR SVM CLASSIFICATION USING A LIMITED TRAINING SET. THE NUMBERS IN BRACKETS ARE THE NUMBERS OF FEATURES FROM THE SPECTRAL DATA AND FROM THE MORPHOLOGICAL DATA, RESPECTIVELY.

	Cum. Var.	Features	OA	AA	κ
Spectral	-	103	71.25	75.79	63.70
EMP	-	27	75.79	80.72	69.00
Spec. EMP	-	130	75.25	81.23	68.66
DBFE	99	66 (44,12)	77.33	83.77	71.31
	95	32 (25,7)	81.34	82.84	75.65
	90	23 (18,5)	74.55	79.15	67.45
	80	16 (12,4)	72.28	80.35	65.25
NWFE	70	11 (9,2)	61.31	72.04	52.26
	99	53 (35,18)	85.42	87.48	80.87
	95	25 (14,11)	81.04	85.42	75.84
	90	17 (10,7)	79.05	83.72	73.27
	80	13 (7,6)	84.77	85.65	80.09
70	9 (5,4)	83.10	84.35	77.84	

very dense urban area.

Regarding the computing time, the results for the training and the classification are reported in Table XIII. As expected, using feature extraction methods reduces the processing time for both the training and the classification.

Classification maps for the different approaches are shown in Fig. 6.

D. Small training set experiment: University Area

To assess the effectiveness of the proposed methodology for a limited training set, we have randomly extracted a few training samples from the training set. For this experiment, we used 20 samples for each class, which represents less than 5% of the original training set. We have used the same EMP but had some problems with the DBFE, the covariance matrix was non-invertible (the NWFE does not suffer from this problem). In order to overcome this shortcoming and to apply the DBFE anyway, we use the *leave on out covariance* (LOOC) to estimate the covariance matrix and perform a *statistical enhancement* with unlabeled samples, both algorithms were implemented in the MultiSpec© software [37], [38]. We have repeated the training samples selection and the classification process five times, and the mean classification results are reported in the paper.

TABLE X
PAVIA CENTER. SUMMARY OF THE GLOBAL AND THE CLASS-SPECIFIC TEST ACCURACIES IN PERCENTAGE FOR SVM CLASSIFICATION. THE NUMBERS OF FEATURES FROM THE SPECTRAL DATA AND THE MORPHOLOGICAL DATA, RESPECTIVELY, ARE GIVEN IN BRACKETS.

	ML (DBFE)	Spectral	EMP	Spec. EMP	DBFE 99%	NWFE 99%
Features	29	102	27	129	66 (51,15)	64 (44,20)
OA	94.50	97.67	98.69	99.69	98.65	98.87
AA	94.00	95.60	97.69	98.07	97.30	97.95
κ	-	96.71	98.15	98.15	98.10	98.41
Class 1	91.50	98.35	99.08	98.66	99.17	99.21
Class 2	92.00	91.23	91.62	93.52	90.00	92.49
Class 3	97.70	96.76	96.18	95.95	96.54	96.76
Class 4	86.90	88.45	98.40	98.77	98.92	99.55
Class 5	95.60	96.97	98.81	99.42	99.27	99.74
Class 6	94.40	96.32	97.98	98.36	98.45	98.70
Class 7	96.40	96.01	97.89	98.22	97.91	98.41
Class 8	99.30	99.40	99.74	99.79	99.81	99.72
Class 9	92.30	99.93	99.44	99.93	98.60	96.93

TABLE XI
PAVIA CENTER. STATISTICAL SIGNIFICANCE OF DIFFERENCES IN CLASSIFICATION ACCURACIES.

	EMP/Spectral	EMP/Spec. EMP	EMP/DBFE 99%	EMP/NWFE 99%	Spec. EMP/DBFE 95%	Spec. EMP/NWFE 80%	DBFE 95%/NWFE 80%
Z	27.84	-0.06	1.44	-8.14	1.42	-7.75	-9.44

TABLE XIII
PAVIA CENTER. PROCESSING TIME IN SECONDS AS FUNCTION OF DIMENSIONALITY AND NUMBER OF SUPPORT VECTORS.

	Spectral	EMP	Spec. EMP	DBFE 95%	NWFE 80%
Dimension	102	27	129	66	66
Training (s)	5178	1569	5909	3520	3962
Number of SVs	691	265	401	558	408
Classification (s)	143	20	105	73	47

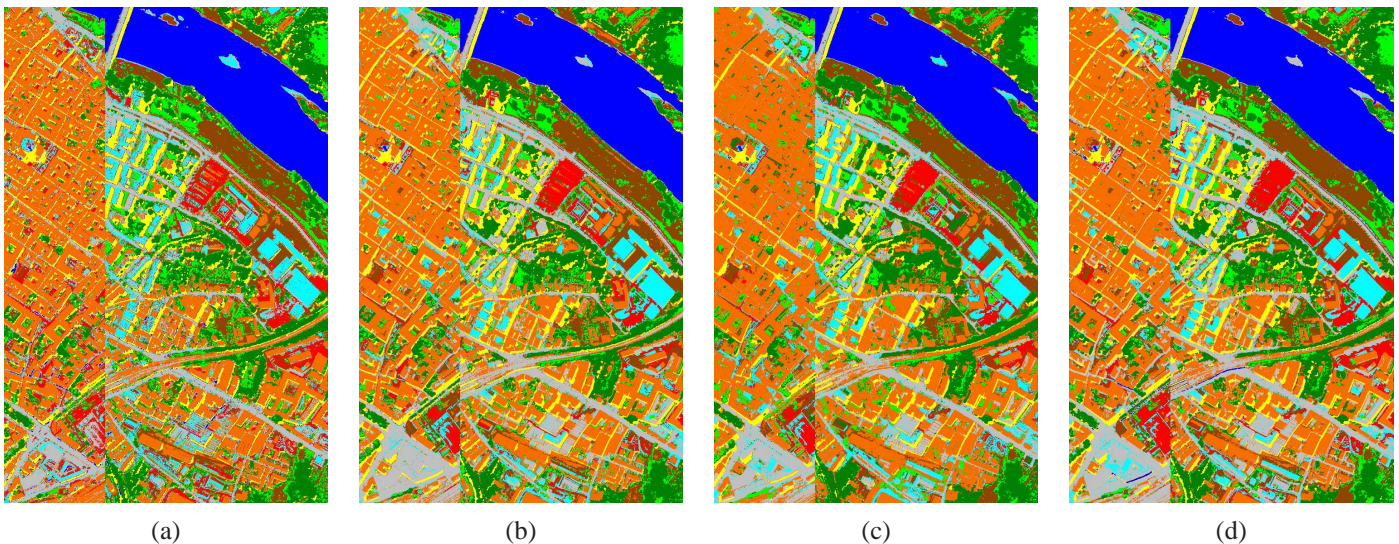


Fig. 6. **Pavia Center:** Classification map obtained with SVM from: (a) the original hyperspectral data, (b) the EMP, (c) 66 DBFE features and (d) 64 NWFE features. Classification accuracies are reported in the Table X.

As with the previous experiments, we perform the classification using the spectral or the morphological feature with SVM. The ML produced very poor results, simply close to random classification and hence not reported. The global accuracies are reported in Table XIV. Statistical significance of differences is reported in Table XV.

First of all, the test results are lower than those in Table V and VII, due to the limited training set. For instance, with the

concatenated feature vector, the overall accuracy and the κ are respectively, 83.53% and 79.13% for the original training set while using a limited training set, the overall accuracy and the κ are respectively, 75.35% and 68.66%. Nevertheless, with a very small training set, the results are still good.

For the feature extraction, NWFE with 99% of the cumulative variance provides the best results: The obtained overall accuracy is 85.42% and the κ is 80.87%, which is closed to the best

results obtained with the full training set (OA=97.87% and $\kappa = 84.40\%$, see Table V). The $|Z|$ between the best results with limited training and the best results with full training set is equal to 13.65.

Furthermore, the accuracies are better than those obtained with the full training set with the spectral or morphological information alone. It is also important to note that NWFE outperforms better DBFE without any statistical enhancement.

Considering the processing time, with only 20 samples for each class, the training as well as the classification of the entire data set are done in 1 or 2 seconds.

VII. CONCLUSION

Classification of hyperspectral data with a fine spatial resolution has been investigated. The contribution of this work is a methodology to include both spatial and spectral information in the classification process by a data fusion scheme. Experimental results on two ROSIS data sets showed excellent accuracies and improvements compared to those obtained with pixel-based classifiers and the EMP-based classifier.

The use of feature extraction was motivated by the fact that the full stacked vector contains a lot of redundancy, because there is a redundancy in the hyperspectral data [3] as well as in the EMP [11], which was confirmed by the experiments. On the other hand, SVM are known to be robust to dimensionality. Therefore, the use of feature reduction for SVM could be disputable. However, in the experiments lower dimensional data decreased the processing time, which can be crucial for some applications, and more important it has been showed that SVM can suffer from the dimensionality when many features are irrelevant [39]. By construction, the stacked vector may contain many copies of the same information and a feature extraction step may finally be needed to ensure correct classification on every data set, which is confirmed by the experiments (Usefulness of features reduction for the classification of remote sensing data with SVM was also assessed in [40]).

It is clear that feature extraction helps for the classification of hyperspectral data but it is not clear which one of the feature extraction methods should be used for the fusion of morphological and spectral features. From a theoretical point of view, the NWFE was derived because of some intrinsic problems with the DBFE [3], i.e., "DBFE can involve lengthly calculations and more significantly it does not perform as well for small numbers of training samples". Hence, the NWFE might be more preferable, especially when a small training set is available. The experiments performed with a limited training set confirmed that.

In conclusion, the proposed fusion method succeed in taking advantage of the spatial and the spectral information simultaneously. It outperformed previous results [2], [10]. Our current research is oriented to the definition of additional spatial features, such as textural characteristics [41], to be include in the feature vectors.

ACKNOWLEDGMENT

The authors appreciate the contributions of Mr. Jon Aevor Palmason to this research. Furthermore, the authors would like

to thank the IAPR - TC7 for providing the data and Prof. Paolo Gamba and Prof. Fabio Dell'Acqua of the University of Pavia, Italy, for providing reference data. The authors also thank the reviewers for their many helpful comments. This research was supported in part by the Research Fund of the University of Iceland and the Jules Verne Program of the French and Icelandic governments (PAI EGIDE).

REFERENCES

- [1] M. Pesaresi and J. A. Benediktsson, "A new approach for the morphological segmentation of high-resolution satellite imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 2, pp. 309–320, Feb. 2001.
- [2] J. A. Palmason, J. A. Benediktsson, J. R. Sveinsson, and J. Chanussot, "Classification of hyperspectral data from urban areas using morphological preprocessing and independent component analysis," in *Proc. IEEE International Geoscience and Remote Sensing Symposium*, vol. 1. IGARSS'05. Proceedings, Jul. 2005, pp. 176–179.
- [3] D. A. Landgrebe, *Signal Theory Methods in Multispectral Remote Sensing*. New Jersey: John Wiley and Sons, 2003.
- [4] V. Vapnik, *Statistical Learning Theory*. Wiley, New York, 1998.
- [5] G. Poggi, G. Scarpa, and J. Zerubia, "Supervised segmentation of remote sensing images based on a tree-structure MRF model," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 8, pp. 1901–1911, Aug. 2005.
- [6] Q. Jackson and D. A. Landgrebe, "Adaptive bayesian contextual classification based on markov random field," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, pp. 2454–2463, Nov. 2002.
- [7] J. A. Benediktsson, M. Pesaresi, and K. Arnason, "Classification and feature extraction for remote sensing images from urban areas based on morphological transformations," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 9, pp. 1940–1949, Sep. 2003.
- [8] F. Dell'Acqua, P. Gamba, A. Ferrari, J. A. Palmason, and J. A. Benediktsson, "Exploiting spectral and spatial information in hyperspectral urban data with high resolution," *IEEE Geosci. Remote Sens. Lett.*, vol. 1, no. 4, pp. 322–326, Oct. 2004.
- [9] M. Fauvel, J. Chanussot, J. A. Benediktsson, and J. R. Sveinsson, "Spectral and spatial classification of hyperspectral data using SVMs and morphological profiles," in *Proc. IEEE International Geoscience and Remote Sensing Symposium*. IGARSS'07. Proceedings, Jul. 2007.
- [10] J. A. Palmason, J. A. Benediktsson, J. R. Sveinsson, and J. Chanussot, "Fusion of morphological and spectral information for classification of hyperspectral urban remote sensing data," in *Proc. IEEE International Geoscience and Remote Sensing Symposium*. IGARSS'06. Proceedings, Jul. 2006.
- [11] J. A. Benediktsson, J. A. Palmason, and J. Sveinsson, "Classification of hyperspectral data from urban areas based on extended morphological profiles," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 3, pp. 480–491, Mar. 2005.
- [12] G. Camps-Valls, L. Gomez-Chova, J. Munoz-Mari, J. Vila-Francés, and J. Calpe-Maravilla, "Composite kernels for hyperspectral image classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 3, no. 1, pp. 93–97, Jan. 2006.
- [13] G. Mercier and F. Girard-Arduin, "Partially supervised oil-slick detection by SAR imagery using kernel expansion," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 10, pp. 2839–2846, Oct. 2006.
- [14] M. Fauvel, J. Chanussot, and J. A. Benediktsson, "A joint spatial and spectral SVM's classification of panchromatic images," in *Proc. IEEE Geoscience and Remote Sensing Symposium*. IGARSS '07. Proceedings, Jul. 2007.
- [15] F. Melgani and L. Bruzzone, "Classification of hyperspectral remote sensing images with support vector machines," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, pp. 1778–1790, Aug. 2004.
- [16] E. Zheng, P. Li, and Z. Song, "Performance analysis and comparison of neural networks and support vectors machines classifier," in *Proceeding of the 5th World Congress on Intelligent Control and Automation*, Jun. 2004.
- [17] G. H. Halldorsson, J. A. Benediktsson, and J. R. Sveinsson, "Support vector machines in multisource classification," in *Proc. IEEE International Geoscience and Remote Sensing Symposium*, vol. 3. IGARSS '03. Proceedings, Jul. 2003, pp. 2054–2056.
- [18] P. Soille and M. Pesaresi, "Advances in mathematical morphology applied to geoscience and remote sensing," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 9, pp. 2042–2055, September 2002.

TABLE XV
UNIVERSITY AREA. STATISTICAL SIGNIFICANCE OF DIFFERENCES IN CLASSIFICATION ACCURACY FOR A LIMITED TRAINING SET.

	EMP/Spectral	EMP/Spec. EMP	EMP/DBFE 95%	EMP/NWFE 99%	NWFE 99%/DBFE 95%
Z	17.5	2.67	-25.4	-51.14	21.32

- [19] P. Soille, *Morphological Image Analysis, Principles and Applications- 2nd edition*. Springer, 2003.
- [20] C. Lee and D. A. Landgrebe, "Feature extraction based on decision boundaries," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 15, no. 4, pp. 388–400, Apr. 1993.
- [21] B.-C. Kuo and D. A. Landgrebe, "A robust classification procedure based on mixture classifiers and nonparametric weighted feature extraction," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 11, pp. 2486–2494, Nov. 2002.
- [22] J. Chanussot, J. A. Benediktsson, and M. Fauvel, "Classification of remote sensing images from urban areas using a fuzzy possibilistic model," *IEEE Geosci. Remote Sens. Lett.*, vol. 3, no. 1, pp. 40–44, January 2006.
- [23] J. A. Gualtieri and R. F. Cromp, "Support vector machines for hyperspectral remote sensing classification," in *Proceeding of the SPIE*, SPIE, Ed., vol. 3584, 1998, pp. 221–232.
- [24] J. A. Gualtieri and S. Chettri, "Support vector machines for classification of hyperspectral data," in *Proc. IEEE International Geoscience and Remote Sensing Symposium*, vol. 2. IGARSS '00. Proceedings, Jul. 2000, pp. 813–815.
- [25] G. M. Foody and A. Mathur, "A relative evaluation of multiclass image classification by support vector machines," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, pp. 1335–1343, Jun. 2004.
- [26] M. Fauvel, J. Chanussot, and J. A. Benediktsson, "Evaluation of kernels for multiclass classification of hyperspectral remote sensing data," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing*, vol. 2. ICASSP'06. Proceedings, May 2006, pp. 813–816.
- [27] L. Bruzzone, M. Chi, and M. Marconcini, "A novel transductive SVM for semisupervised classification of remote-sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 11, pp. 3363–3373, Nov. 2006.
- [28] M. Chi and L. Bruzzone, "Semisupervised classification of hyperspectral images by SVMs optimized in the primal," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 6, pp. 1870–1880, Jun. 2007.
- [29] C. Burges, "A tutorial on support vector machines for pattern recognition," *Data Mining and Knowledge Discovery*, vol. 2, no. 2, pp. 121–167, 1998.
- [30] N. Aronszajn, "Theory of reproducing kernel," Harvard University, Division of engineering sciences, Tech. Rep. 11, 1950.
- [31] C.-W. Hsu and C.-J. Lin, "A comparison of methods for multiclass support vector machines," *IEEE Trans. Neural Netw.*, vol. 13, pp. 415–425, Mar. 2002.
- [32] C. Sanchez-Hernandez, D. S. Boydand, and G. M. Foody, "One-class classification for mapping a specific land-cover class: SVDD classification of fenland," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 4, pp. 1061–1073, Apr. 2007.
- [33] C.-C. Chang and C.-J. Lin, *LIBSVM: a library for support vector machines*, 2001, software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.
- [34] M. Fauvel, J. Chanussot, and J. A. Benediktsson, "Decision fusion for the classification of urban remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 10, Oct. 2006.
- [35] J. A. Richards and X. Jia, *Remote Sensing Digital Image Analysis: An Introduction*. Springer, 1999.
- [36] G. M. Foody, "Thematic map comparison: Evaluating the statistical significance of differences in classification accuracy," *Photogrammetric Engineering & Remote Sensing*, vol. 70, no. 5, pp. 627–633, May 2004.
- [37] J. Hoffbeck and D. A. Landgrebe, "Covariance matrix estimation and classification with limited data," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 18, no. 7, pp. 763–767, Jul. 1996.
- [38] B. M. Shahshahani and D. A. Landgrebe, "The effect of unlabeled samples in reducing the small sample size problem and mitigating the Hughes phenomenon," *IEEE Trans. Geosci. Remote Sens.*, vol. 32, no. 5, pp. 1087–1095, Sep. 1994.
- [39] O. Chapelle, "Support vector machines: Induction principle, adaptive tuning and prior knowledge," Ph.D. dissertation, Université Pierre et Marie Curie, <http://www.kyb.mpg.de/publication.html?publ=2167>, 2002.
- [40] Y. Bazi and F. Melgani, "Toward an optimal SVM classification system for hyperspectral remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 11, pp. 3374–3385, Nov. 2006.
- [41] I. Epifanio and P. Soille, "Morphological texture features for unsupervised and supervised segmentations of natural landscapes," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 4, pp. 1074–1083, Apr. 2007.



Mathieu Fauvel (S'06) graduated in electrical engineering from the Grenoble National Polytechnical Institute (INPG), Grenoble, France, in 2004. He received the M.S. and Ph.D. degrees in image and signal processing from the INPG in 2004 and 2007, respectively. Dr. Fauvel is currently working at the GIPSA-lab.

His research interests are remote sensing, data fusion, pattern recognition, multicomponent signal and image processing. He is serving as a reviewer for the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING and the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS.

MOTE SENSING and the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS.



Jón Atli Benediktsson (S'84-M'90-SM'99-F'04) received the Cand.Sci. degree in electrical engineering from the University of Iceland, Reykjavik, in 1984, and the M.S.E.E. and Ph.D. degrees from Purdue University, West Lafayette, IN, in 1987 and 1990, respectively. He is currently Director of Academic Development and Innovation and Professor of Electrical and Computer Engineering at the University of Iceland. He has held visiting positions at the Department of Information and Communication Technology, University of Trento, Trento, Italy (2002-present), School

of Computing and Information Systems, Kingston University, Kingston upon Thames, U.K. (1999-2004), the Joint Research Centre of the European Commission, Ispra, Italy (1998), Denmark's Technical University (DTU), Lyngby (1998), and the School of Electrical and Computer Engineering, Purdue University (1995). He was a Fellow at the Australian Defence Force Academy, Canberra, A.C.T., Australia, in August 1997. From 1999 to 2004, he was a Chairman of the energy company Metan Ltd. His research interests are in remote sensing, pattern recognition, neural networks, image processing, and signal processing, and he has published extensively in those fields. Dr. Benediktsson is Editor of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING (TGARS) and Associate Editor of the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS. He was Associate Editor of TGARS from 1999 to 2002. He co-edited (with Prof. D. A. Landgrebe) a Special Issue on Data Fusion of TGARS (May 1999). In 2002, he was appointed Vice President of Technical Activities in the Administrative Committee of the IEEE Geoscience and Remote Sensing Society (GRSS) and (with P. Gamba and G. G. Wilkinson) a Special Issue on Urban Remote Sensing from Satellite (October 2003). From 1996 to 1999, he was the Chairman of GRSS Technical Committee on Data Fusion and was elected to the Administrative Committee of the GRSS for the term 2000 to 2002, and in 2002, he was appointed Vice President of Technical Activities of GRSS. He was the founding Chairman of the IEEE Iceland Section and served as its Chairman from 2000 to 2003. Currently, he is the Chairman of the University of Iceland's Quality Assurance Committee (2006-). He was the Chairman of the University of Iceland's Science and Research Committee (1999-2005), a member of Iceland's Science and Technology Council (2003-2006), and a member of the Nordic Research Policy Council (2004). He was a member of a NATO Advisory Panel of the Physical and Engineering Science and Technology SubProgramme (2002-2003). Dr. Benediktsson is a Fellow of the IEEE. He received the Stevan J. Kristof Award from Purdue University in 1991 as outstanding graduate student in remote sensing. In 1997, he was the recipient of the Icelandic Research Council's Outstanding Young Researcher Award, in 2000, he was granted the IEEE Third Millennium Medal, in 2004, he was a co-recipient of the University of Iceland's Technology Innovation Award, in 2006 he received the yearly research award from the Engineering Research Institute of the University of Iceland, and in 2007, he received the Outstanding Service Award from the IEEE Geoscience and Remote Sensing Society. He is a member of Societas Scinetiarum Islandica and Tau Beta Pi.



Johannes R. Sveinsson (S'86-M'90-SM-02) received the B.S. degree from the University of Iceland, Reykjavik, and the M.S. (Eng.) and Ph.D. degrees from Queen's University, Kingston, ON, Canada, all in electrical engineering. He is currently a Professor in the Department of Electrical and Computer Engineering, University of Iceland. From 1981 to 1982, he was with the Laboratory of Information Technology and Signal Processing, University of Iceland. He was a Visiting Research Student with the Imperial College of Science and Technology, London, U.K. from 1985 to 1986. At

Queen's University, he held teaching and research assistantships and received Queen's Graduate Awards. From November 1991 to 1998, he was with the Engineering Research Institute, University of Iceland, as a Senior Member of the research staff and a Lecturer in the Department of Electrical Engineering, University of Iceland. His current research interests are in systems and signal theory.



Jocelyn Chanussot (SM'04) received the M.Sc. degree in electrical engineering from the Grenoble Institute of Technology (Grenoble INP), France, in 1995, and the Ph.D. degree from the Savoie University, Annecy, France, in 1998. In 1999, he was with the Geography Imagery Perception laboratory (GIP) for the Delegation Generale de l'Armement (DGA-French National Defense Department). Since 1999, he has been with Grenoble INP as an Assistant Professor from 1999 to 2005), an Associate Professor from 2005 to 2007, and currently a Professor of signal and image

processing. He is conducting his research at the Grenoble Images Speech Signals and Automatics Lab (GIPSA-Lab, Grenoble, France). His research interests include image analysis, multicomponent image processing, nonlinear filtering and data fusion in remote sensing. Prof. Chanussot is an Associate Editor for the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING and for Pattern Recognition. He is the Co-Chair of the GRS Data fusion Technical Committee and a Member of the Machine Learning for Signal Processing Technical Committee of the IEEE Signal Processing Society. He is the founding president of IEEE Geoscience and Remote Sensing French chapter.