Decision Fusion for the Classification of Hyperspectral Data: Outcome of the 2008 GRS-S Data Fusion Contest

Giorgio Licciardi, Fabio Pacifici, *Student Member, IEEE*, Devis Tuia, *Student Member, IEEE*, Saurabh Prasad, *Member, IEEE*, Terrance West, *Student Member, IEEE*, Ferdinando Giacco, Christian Thiel, Jordi Inglada, Emmanuel Christophe, Jocelyn Chanussot, *Senior Member, IEEE*, and Paolo Gamba, *Senior Member, IEEE*

***Abstract*—The 2008 Data Fusion Contest organized by the IEEE Geoscience and Remote Sensing Data Fusion Technical Committee deals with the classification of high-resolution hyperspectral data from an urban area. Unlike in the previous issues of the contest, the goal was not only to identify the best algorithm but also to provide a collaborative effort: The decision fusion of the best individual algorithms was aiming at further improving the classification performances, and the best algorithms were ranked according to their relative contribution to the decision fusion. This paper presents the five awarded algorithms and the conclusions of the contest, stressing the importance of decision fusion, dimension reduction, and supervised classification methods, such as neural networks and support vector machines.**

***Index Terms*—Classification, decision fusion, hyperspectral imagery.**

I. Introduction

THE DATA Fusion Contest has been organized by the Data Fusion Technical Committee (DFTC) of the IEEE Geoscience and Remote Sensing Society and has been annually proposed since 2006. It is a contest open not only to DFTC members but also to everyone. The aim of the Data Fusion Contest is to evaluate existing methodologies at the research or operational level to solve remote sensing problems using data from different sensors. The main aim of this contest is to provide a benchmark to the researchers interested in a class of data fusion problems, starting with a contest and then allowing the data and results to be used as reference for the widest community, inside and outside the DFTC. The first issue of the contest was devoted to pansharpening [1]. In 2007, the contest was related to urban mapping using radar and optical data [2].

In 2008, the contest was dedicated to the classification of very high resolution hyperspectral data. A hyperspectral data set was distributed to every participant, and the task was to obtain a classified map as accurate as possible with respect to the ground truth data, depicting land-cover and land-use classes. The ground truth was kept secret, but training pixels could be selected by the participants by photointerpretation in order to apply supervised methods. The data set consisted of airborne data from the reflective optics system imaging spectrometer (ROSIS-03) optical sensor. The flight over the city of Pavia, Italy, was operated by the Deutschen Zentrum fur Luft-und Raumfahrt (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 0.86 μm. Thirteen noisy bands have been removed. The dimension of the distributed data set is hence 102. The spatial resolution is 1.3 m per pixel. For the contest, five classes of interest were considered, namely, buildings, roads, shadows, vegetation, and water. Everyone could enter the contest and download the data set. After classification, the participant could upload the resulting map for an automatic evaluation of the classification performances (confusion matrix and average accuracy). The participating teams were allowed to upload as many different results as they wished.

At any given time, the five best maps were combined using majority voting (MV) and reranked according to their respective contribution to the fused result. The best seven individual algorithms were listed in real time on the data fusion contest website (http://tlclab.unipv.it/dftc/home.do), together with the result of the fusion. Please note that the website is still open and everyone can use it as a benchmark to test any new algorithm.

The contest was open for three months. At the end of the contest, 21 teams had uploaded over 2100 classification maps! A closer look reveals that one single team actually submitted over 1200 results (but we should underline that it did not rank in the top five teams), while the other 1000 entries are spread over the remaining 20 teams. The five best individual classification maps have been fused together. The final corresponding teams have been awarded with an IEEE Certificate of Recognition during the Chapters and Technical Committees' Dinner at the IEEE

Manuscript received October 29, 2008; revised April 1, 2009. First published October 13, 2009; current version published October 28, 2009.

G. Licciardi and F. Pacifici are with the Earth Observation Laboratory, Tor Vergata University, 00133 Rome, Italy.

D. Tuia is with the University of Lausanne, 1015 Lausanne, Switzerland.

S. Prasad and T. West are with the Mississippi State University, Starkville, MS 39762 USA.

F. Giacco is with the Department of Physics, University of Salerno, 84084 Salerno, Italy.

C. Thiel is with the University of Ulm, 89069 Ulm, Germany.

J. Inglada is with the Centre National d'Etudes Spatiales, 31401 Toulouse, France.

E. Christophe is with the Centre for Remote Imaging, Sensing and Processing, National University of Singapore, Singapore 119260.

J. Chanussot is with the Laboratoire Grenoblois de l'Image, de la Parole, du Signal et de l'Automatique, Grenoble Institute of Technology, 38402 Grenoble, France.

P. Gamba is with the University of Pavia, 27100 Pavia, Italy.

Digital Object Identifier 10.1109/TGRS.2009.2029340

International Geoscience and Remote sensing Symposium in Boston in July 2008.

The remainder of this paper is organized as follows. First, the best five algorithms are detailed.

1) Section II presents the work of Giorgio Licciardi and Fabio Pacifici. They use different standard classifiers [three neural networks (NNs) and two maximum likelihood (ML) classifiers] and perform an MV between different outputs.
2) Section III presents the work of Devis Tuia and Frederic Ratle. They use both spectral and spatial features. The spectral features are a six-principal-component (PC) analysis (PCA) extraction of the initial pixel's vector value. The spatial information is extracted using morphological operators. These features are classified by combining several support vector machines (SVM) using MV.
3) Section IV presents the work of Saurabh Prasad and Terrance West.[1] They use wavelet-based preprocessing of the initial spectra followed by a linear discriminant analysis (LDA) and an ML classifier.
4) Section V presents the work of Ferdinando Giacco and Christian Thiel. They use a PCA to reduce the dimension of the data. Spatial information is taken into account with some textural features. The classification is achieved using SVM one-versus-one classifiers, and a spatial regularization is performed on the classification map to eliminate isolated pixels.
5) Section VI presents the work of Jordi Inglada and Emmanuel Christophe. They perform a Bayesian fusion of different classifiers (such as SVM classifiers). The weight assigned to each classifier is determined by the quantitative results it obtained. All these algorithms are available with the ORFEO Toolbox, which is an open source library of image processing algorithms for remote sensing applications (http://www.orfeo-toolbox.org).

Finally, the decision fusion is considered in Section VII, and the conclusions and perspectives drawn by this contest are presented and discussed in Section VIII.

II. MV Between NN and ML Classifiers

A. Reduction of Data Dimensionality

The analysis of hyperspectral imagery usually implicates the reduction of data set dimensionality to decrease the complexity of the classifier and the computational time required with the aim of preserving most of the relevant information of the original data according to some optimal or suboptimal criteria [3], [4]. The preprocessing procedure exploited in this section divides the hyperspectral signatures into adjacent regions of the spectrum and approximates their values by piecewise constant functions. In [5], the authors reduced effectively the input space using the averages of contiguous spectral bands applying piecewise constant functions instead of higher order polynomials. This simple representation has shown to outperform most of the feature reduction methods proposed in the literature, such as PC transform, sequential forward selection, or decision boundary feature extraction [6].

TABLE I
RESULTING SUBBANDS

	Sensor bands		Wavelength (μm)	
	from	*to*	*from*	*to*
B1	1	15	430	486
B2	16	35	490	566
B3	36	65	570	686
B4	66	75	690	726
B5	78	82	730	766
B6	86	90	770	786
B7	91	95	790	834

TABLE II
TRAINING SAMPLES USED FOR THE SUPERVISED CLASSIFIERS

	Buildings	Roads	Shadows	Vegetation	Water
Set 1	132,369	18,914	20,356	53,065	43,104
Set 2	33,168	6,525	3,260	14,323	26,816
Set 3	45,268	5,210	1,524	17,485	20,367

Assume S_{ij} to be the value of the ith pixel in the jth band, with a total of N pixels. The spectral signatures of each class extracted from ground truth pixels have been partitioned into a fixed number of contiguous intervals with constant intensities minimizing the mean-square error

$$H = \sum_{k=1}^{K} \sum_{i=1}^{N} \sum_{j \in I_k} (S_{ij} - \mu_{ik})^2 \quad (1)$$

where a set of K breakpoints defines continuous intervals I_k, while μ_{ik} represents the mean value of each pixel's interval between breakpoints. A number of $K = 7$ breakpoints were found to be a reasonable compromise between model complexity and computational time, and the resulting partitions are reported in Table I.

B. Classification Phase

In the literature, NNs and SVMs have been widely used since they do not require any specific probabilistic assumptions of the class distribution, in opposition to parametric classifiers, such as ML. The classifier scheme exploited here is a combination of single decision maps. In [7], it has been demonstrated that combining the decisions of independent classifiers can lead to better classification accuracies. The combination can be implemented using a variety of strategies, among which MV is the simplest, and it has been found to be as effective as more complicated schemes [8], [9].

MV was used here on five independent maps resulting from two different methods, i.e., three NNs and two ML classifiers. For each method, the input space was composed by the seven features obtained by reducing the sensor bands, while the outputs were the five classes of interest. For training the supervised classifiers, we have defined three different training sets, varying the number of samples, as reported in Table II. In the following, we briefly recall the classification methods and the setting used.

1) NNs: The topology of a multilayer perceptron network [10] has been determined through an optimization of the number of hidden layers and units, based on the results reported in the literature, on previous experiences and on a specific numerical analysis [11]. Two hidden layers have been found to be a

[1]The authors would like to acknowledge the active participation of Jeff Brantley, Jacob Bowen, and Matthew Lee to this work. They are all with the Mississippi State University.

TABLE III
TRAINING SET CLASSIFICATION ACCURACIES FOR NN, ML, AND MV

	NN 1 (set 1)	NN 2 (set 2)	NN 3 (set 3)	ML 1 (set 1)	ML 2 (set 2)	MV
Acc. (%)	95.6	95.4	95.1	95.0	94.9	96.3
K-coef.	0.936	0.932	0.929	0.927	0.925	0.946

TABLE IV
CONFUSION MATRIX; TRUE CLASSES GIVEN BY ROWS

Class	1	2	3	4	5
99.65%	213359	391	155	203	0
97.05%	246	10430	0	71	0
98.95%	143	27	16245	2	0
99.97%	2	5	1	24480	0
100.00%	0	0	0	0	10961
	99.82%	96.10%	99.05%	98.89%	100%

suitable choice, while the number of hidden neurons was found using a growing method, progressively increasing the number of elements. The variance of the classification accuracy for different initializations of the weights was computed to monitor the stability of the topology. The configuration 7-25-25-5 maximized the accuracy and minimized the instability of the results. Successively, three independent NNs were trained with sets 1, 2, and 3 (see Table II), providing three different maps.

2) ML: ML is a well-known parametric classifier, which relies on the second-order statistics of a Gaussian probability density function for the distribution of the feature vector of each class. ML is often used as a reference for classifier comparisons because it represents an optimal classifier in the case of normally distributed class probability density functions [12]. ML classification was performed using sets 1 and 2 (see Table II), providing two different maps.

The results from the five classification maps were combined using MV to obtain the final map. The algorithm of MV was implemented by following two simple rules.

1) A class is the winner if it is recognized by the majority of the classifiers.
2) In the case of balance voting, the winner class is the one with the highest Kappa (K) coefficient.

The improvement derived from MV is reported in Table III, where the K-coefficients (based on training sets) obtained from five classifications are compared with the one of the final result.

Table IV presents the corresponding final confusion matrix. The score is 0.9884.

III. MORPHOLOGICAL FEATURES AND SVM CLASSIFIER

The proposed method uses both spectral and spatial information to train an SVM classifier. A brief description of the input features and of the classifier exploited is discussed in this paragraph.

A. Spectral and Spatial Features

The PCA was used to extract spectral information from the original image. Specifically, the six first PCs have been retained for analysis, as shown from the component composition in Fig. 1(b). These features count for 99.9% of the variance contained in the original hyperspectral bands.

(a) (b)

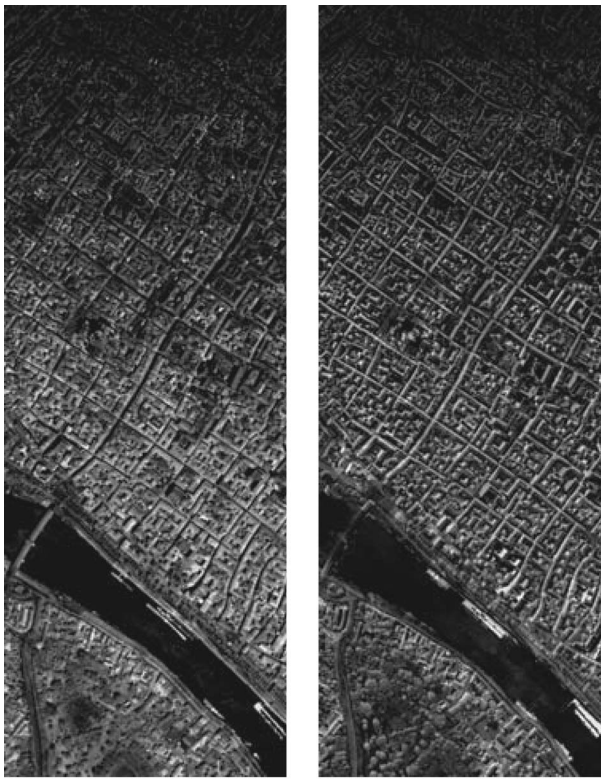
Fig. 1. (a) First PC. (b) Six PCs retained.

Morphological operators [13], [14] have been added to include information about the spatial neighborhood of the pixels. Mathematical morphology is a collection of filters called operators based on set theory. Morphological operators have been used in remote sensing to extract information about the shape and structure of the objects in both optical [15], [16] and, more recently, hyperspectral imageries [17]–[19].

An operator is applied using two ensembles: The first is the image to filter g, and the second is a set of known size and shape called the structuring element B. In our setting, and as suggested in [17] and [18], the first PC [shown in Fig. 1(a)] has been used for the extraction of the morphological features. Specifically, top-hat features have been considered. These features are constructed using the three-stage filtering described as follows.

1) Erosion and dilation. For a given pixel on the input image g, erosion $\epsilon_B(g)$ is the pointwise minimum $\wedge$ between all the values of g defined by B when centered on the pixel considered. On the contrary, dilation $\delta_B(g)$ is the pointwise maximum $\vee$ between these same values.
2) Opening and closing. Opening $\gamma_B(g)$ is the dilation of an eroded image and is widely used to isolate brighter (compared to surrounding features) structures in gray-scale images. On the contrary, closing $\phi_B(g)$ is the erosion of a dilated image and allows one to isolate darker structures [20]. The formulation of opening and closing operators is given by

$$\gamma_B(g) = \delta_B\left[\epsilon_B(g)\right] \qquad \phi_B(g) = \epsilon_B\left[\delta_B(g)\right]. \quad (2)$$



(a) (b)

Fig. 2. (a) Opening and (b) closing top-hat features extracted for the Pavia image. The size of the structuring element is increased from top (three pixels) to the bottom (29 pixels) of the images.

TABLE V
LABELED PIXELS FOR THE PAVIA IMAGE

Class	Labeled pixels	Training	Validation	Test
Buildings	84305	13000	12484	58821
Roads	17495	7000	1840	8655
Shadow	11375	7000	758	3617
Vegetation	49730	5000	7770	36960
Water	43104	2000	7148	33956
Total	206009	34000	30000	142009

3) Top hat. Top-hat operators are the residuals of an opening (or a closing) image, when compared to the original image, as

$$TH = g - I(g). \tag{3}$$

If $I = \gamma_B(g)$, the operator is an opening top hat and highlights bright peaks of the image. On the contrary, if $I = \phi_B(g)$, the operator is closing top hat and emphasizes dark peaks of the image, as shown in Fig. 2.

B. Experimental Setup

A total of 206 009 labeled pixels has been identified by careful visual inspection of the hyperspectral image. These samples have been divided into a training set of about 34 000 pixels, a validation set for model selection (about 30 000 pixels), and a test containing the remaining 142 009 pixels, as shown in Table V.

As discussed previously, the input space takes both spectral and spatial features into account. The six first PCs have been used as spectral information, while 28 spatial features have been extracted by applying opening and closing top-hat operators to the first PC using diamond-shaped structuring element with increasing diameter size (from 3 to 29 pixels).

TABLE VI
CONFUSION MATRIX; TRUE CLASSES GIVEN BY ROWS

Class	1	2	3	4	5
99.65%	213351	385	260	107	5
95.80%	414	10296	12	25	0
98.42%	223	35	16158	1	0
99.76%	52	5	1	24430	0
100.00%	0	0	0	0	10961
	99.68%	96.04%	98.34%	99.46%	99.95%

Each feature has been converted to standard scores and stacked in a single 34-D input vector. The classifier is a one-against-all SVM implemented using the Torch 3 library [21]. A radial basis function (RBF) kernel has been used. Model selection has been performed by grid search to find the optimal kernel parameters σ and C.

C. MV of the Best Classification Maps

During the contest, several maps have been uploaded, accounting for different training sets and optimal kernel parameters. Eventually, each classification map improving the previous solution has been combined using MV: A pixel received the label of the class assigned by most of the models. In the case where no class prevailed, the pixels receive the label of the map showing the highest Kappa coefficient.

Table VI presents the corresponding final confusion matrix. The score is 0.9858.

IV. GROUND-COVER MAPPING USING SUPERVISED CLASSIFICATION AND MORPHOLOGICAL PROCESSING

In this approach, we employ discrete wavelet transform (DWT)-based processing of the hyperspectral signatures, followed by LDA transformation and pixelwise ML classification for creating a ground-cover map of the satellite imagery. The LDA transformation and ML classifiers are trained using the training data extracted from the regions of interest provided to all contest participants. The resulting ground-cover map is then postprocessed by an appropriate morphological operation to minimize the salt-and-pepper classification noise introduced because of the use of pixelwise (per-pixel) classification. The DWT-based preprocessing of the hyperspectral signatures provides a multiresolution information representation. The mother wavelet employed in this approach is the Daubechies wavelet (implemented using the Daubechies 9/7 filter bank), which resulted in a feature vector comprising of DWT coefficients per pixel. Data from this high-dimensional space were projected onto a reduced-dimensional space by employing the LDA algorithm. LDA seeks to find a linear transformation, such that the within-class scatter is minimized and the between-class scatter is maximized. The transformation is determined by maximizing Fisher's ratio, which can be solved as a generalized eigenvalue problem.

The between- and within-class scatter matrices are learned from the training data. Since it is designed to maximize class separation in the projected space, LDA is an appropriate dimensionality reduction approach for the land-cover classification task at hand.

TABLE VII
CONFUSION MATRIX; TRUE CLASSES GIVEN BY ROWS

Class	1	2	3	4	5
99.42%	212873	237	937	61	0
94.06%	374	10109	13	251	0
96.78%	252	160	15889	78	38
98.98%	108	92	49	24238	1
100.00%	0	0	0	0	10961
	99.66%	95.39%	94.08%	98.42%	99.65%

After performing an LDA transformation on the training and test data, an ML classifier is employed for classifying pixels in the image, which assumes Gaussian class distributions for each class. We assume equal priors for each class. The class membership function for such a classifier is given in [22]. A conventional single-classifier system was sufficient for the given task because the amount of available ground truth was sufficient relative to the feature space dimensionality. Had we had an insufficiently small ground-truth data set for the classification task, the recently developed multiclassifier and decision fusion framework could have been employed for this task [23]. The feature extraction, optimization, and classification approach outlined earlier helps in generating an initial ground-cover map. In order to remove the salt-and-pepper classification noise from this map, morphological postprocessing is performed over it. For each class i, a binary map is created with class i having the label 1 and all other classes having the label 0. A one-pixel dilation is then applied to each set of clustered pixels in the binary map. This dilated mask is then subtracted from the clustered pixels in the binary map which produces a cluster ring. For a cluster smaller than a predetermined class cluster threshold, the cluster ring is placed in the original image, and the class with the largest sum of label pixels in the ring defines the label of the cluster. This is done for all classes. This operation ensures that stray mislabeling of classes (e.g., a building pixel in the middle of a river body) is corrected.

The normalized difference vegetation index (NDVI) is a very good indicator of vegetation in remote sensing applications. As the final postprocessing, we estimated the NDVI value for each pixel in the image. This NDVI map is used to replace the class labels of all nonvegetation pixels in the classification map with vegetation pixels if the corresponding NDVI was high. This ensures that any missed pixels of vegetation pixels using the standard classification approach are identified and corrected. It is worth mentioning that, although we have performed the per-pixel classification in the wavelet domain, we obtained very similar recognition performance (measured by the accuracy) when we performed the classification in the raw reflectance domain. The improvement in the overall classification by introducing wavelet-based processing was marginal.

Table VII presents the corresponding final confusion matrix. The score is 0.9753.

V. THREE-STAGE CLASSIFICATION BASED ON ONE-VERSUS-ONE SVMS

The proposed method is made up of three classification stages with special attention to preprocessing and spatial feature extraction.

1) Preprocessing and Feature Extraction: A PCA of the 102 ROSIS spectral bands is computed. The 26 bands with the most significant PCs are used as spectral input features for the classifier. In addition, we introduced some spatial information extracted from the ROSIS data set: standard deviation calculated on the first PC and on the near infrared/red ratio (bands 102/66), known in remote sensing literature as a way to emphasize the vegetation. We also computed the so-called energy measure, extracted from the well-known gray-level co-occurrence matrix (GLCM), widely used in land-cover mapping [24]. Starting from a pixel in a given position, the GLCM provides a measure of the probability of occurrence of two gray levels separated by a given distance in a given direction (among the horizontal, vertical, left diagonal, and right diagonal). The energy measure is computed, i.e., the summation of squared elements in the GLCM, and the four directions are averaged to remove directional effects; this last choice is due to the absence of a preferred direction in the geometry of the investigated land-cover classes.

Each textural measure is computed on a moving window of 3×3 pixels. The total number of features for the first stage is 29. We worked on a total number of 2133 labeled samples to train the SVMs, which were split into two subsets for training (882) and test (1241) during the parameter optimization phase.

In our second classification stage, in order to improve the discrimination between buildings and streets, we added four new features obtained from the HYPERUSP algorithm. This procedure (implemented in the geographic information system software IDRISI, Andes edition) first makes use of an unsupervised stage in which a prearranged number of hyperspectral signatures are identified looking at the whole ROSIS spectral data set. Then, every pixel of the image is considered as a combination of all the components represented in the signatures computed in the first stage. The coefficients of the four most representative components of the hyperspectral decomposition were selected, adding up to a total number of 33 features for the second classification stage. In addition, the new class "gray building" was introduced, summing up to 1614 labeled pixels for this stage.

2) Classification:

1) First stage. An SVM was used as a multiclass classifier, in a one-versus-one architecture with linear kernel ($C = 1$; RBF or polynomial ones performed not as good), where an SVM_i is built for each possible pair of classes. Presented with a new sample x, each SVM_i answers with the distance $d_i(x)$ that this sample has to its hyperplane. These distances will be converted to probabilities using a sigmoid function [25] with fixed parameters. To incorporate information about class-pair dependences, we proposed [26] to not simply sum up the values per class but use an algorithm based on the statistical Bradley–Terry model. After an iterative process, it produces probabilities that are very plausible given all pairs of classwise comparisons.
2) Second stage. We only looked at those samples that were classified as buildings or streets (class 1 or 2), according to the answers of the previous stage. A one-versus-one SVM with a linear kernel (as described before) was used. This second step increases the overall accuracy from 96.05% to 96.41%.
3) Third stage. A simple filter was used to avoid lonely pixels which are classified differently from their neighbors.

TABLE VIII
CONFUSION MATRIX; TRUE CLASSES GIVEN BY ROWS

Class	1	2	3	4	5
98.92%	211804	1906	262	136	0
86.79%	1403	9327	6	11	0
99.32%	29	57	16306	1	24
99.87%	31	0	1	24456	0
100.00%	0	0	0	0	10961
	99.31%	82.61%	98.38%	99.40%	99.78%

Considering a window of 3×3 surrounding a selected pixel, if the majority of the pixels belong to the same class, the central pixel is assigned to it.

Looking at the confusion matrix (see Table VIII; the score is 0.9641) and also the final map, one can observe that there are very few errors, except for the classes 1 and 2 (buildings and streets). By visual inspection of a natural color composition of ROSIS bands, we found that our classification procedure had still some difficulties in telling gray roofs from streets. Red roofs were classified correctly.

The rather powerful SVM with Bradley–Terry coupled output outperformed some other classifiers tested, and the second stage we implemented proved to alleviate the street/building problem. For even better results, we think that more structural features would be needed.

VI. BAYESIAN FUSION

Given the fact that different classifiers have different performances for different kinds of classes, it was interesting to perform some classifier fusion. Several classification strategies with different refinements were defined to improve the shortcoming we notice during the first tentatives. The idea is to define several methods, each with its own strengths and weaknesses, and to combine the results. We implemented several SVM classifiers using different input features and training sets and applied Bayesian fusion with two approaches.

The first point that we noticed was that SVM classifiers were found to be very sensitive to the training sets. As no training set was provided for the challenge, several training sets were created with different characteristics: including border pixels or not, exhaustive classification of small areas, etc. Another question was raised concerning the definition of classes: Is inner courtyard considered as road or building? Several training sets were created with these different strategies in mind. Finally, four training sets were used.

The second point concerned the input data: Data provided to the SVM are particularly critical. The first possibility is to use the original image. However, using many bands does not allow one to efficiently differentiate classes; thus, the learning stage is usually very costly as the SVM has to find out the significant information. For hyperspectral data, several preprocessing steps are widely used to reduce data dimensionality. PCA was used to concentrate the information on the first few spectral bands and the eight bands with most energy were kept for the SVM. Similar processing was done for the maximum noise fraction (MNF) keeping the first eight bands. As the SVM is able to classify data even when some features of the feature vector are irrelevant or redundant, both PCA and MNF were also combined.

One shortcoming of the SVM classification is that it is based only on one pixel at a time. Pixels on the edge of classes are usually composed of several classes, and it is particularly difficult to classify these pixels without looking at their environment; most classification errors come from these pixels. The simplest way to introduce a relationship between these pixels was to use a Markov random field to regularize the final classification. A simple Potts model was introduced to reduce the noise on these edge pixels. Such regularization usually increases the final score by 2% in average.

An alternative to this regularization was to apply a blur (mean filter) to the input data. Such blur usually reduces the differences between pixels within one class, thus greatly speeding the learning step, without a significant impact on false classification.

All these data sets, training sets, and classification options led to different classification results. Given the fact that the confusion matrix was computed on about one quarter of the pixels, the idea was then to improve the overall results using performances on this pixel subset. This really corresponds to a real case where a ground truth is available for a portion of the image, and the automatic classification is used to speed up the process without any more human intervention. Several approaches were designed to combine these results.

The first approach consisted in performing the ML fusion (MLF) of different M classifiers using the confusion matrix obtained for each of them. Thus, for a given pixel x_i and for each class $C_k = 1, \ldots, N$, we compute likelihood

$$L(x_i, C_k) = \sum_{j=1}^{M} U_j^k \cdot A_{jk} \tag{4}$$

where U_j^k is a binary valued function which is equal to one if classifier j gives class k and zero if otherwise, and A_{jk} is the diagonal term of the confusion matrix of classifier j for class k.

The MLF consists in taking class k, which maximizes the likelihood for each pixel.

The second approach consisted in performing maximum *a posteriori* fusion, which is actually like MLF, but using the prior probabilities of the different classes $P(k)$

$$L(x_i, C_k) = \sum_{j=1}^{M} U_j^k \cdot A_{jk} \cdot P(k). \tag{5}$$

$P(k)$ can easily be obtained from the output of each classifier, since these are good enough to assume that the proportions of the classes are correct. One can also obtain these proportions by computing a weighted average of the proportions of each classifier. The weights can proportional to the kappa coefficient of each classifier.

Combining several classifications leads to improved results: 1% over the best classification. This result is also more robust as it does not need any fine tuning of the SVM parameters: The worst results will be discarded during the fusion process.

Table IX presents the corresponding final confusion matrix. The score is 0.9612.

VII. DECISION FUSION

The decision fusion of the five best individual results (described in the previous sections) was achieved using a simple majority vote. Table X presents the corresponding final confusion matrix. The score is 0.9921. Even though the final score is less than 1% higher than the best algorithm, it remains the best.

TABLE IX
CONFUSION MATRIX; TRUE CLASSES GIVEN BY ROWS

Class	1	2	3	4	5
98.45%	210786	1426	1881	13	2
97.06%	135	10431	181	0	0
97.27%	88	178	15968	0	183
99.38%	11	139	3	24335	0
100.00%	0	0	0	0	10961
	99.89%	85.68%	85.55%	99.95%	98.34%

TABLE X
CONFUSION MATRIX; TRUE CLASSES GIVEN BY ROWS

Class	1	2	3	4	5
99.76%	213600	229	248	31	0
98.06%	199	10539	2	7	0
99.29%	71	43	16301	1	1
99.93%	8	9	1	24470	0
100.00%	0	0	0	0	10961
	99.87%	97.40%	98.48%	99.84%	99.99%

As a conclusion, one can clearly state that decision fusion is indeed a promising way in order to actually solve the problem of classification in hyperspectral imagery. One can think of the result of this contest as the "metaclassifier" everyone has been dreaming of, but no one ever did implement such a classifier.

As a matter of fact, it requires the perfect mastering, implementation, and tuning of very different up-to-date techniques, from dimension reduction to feature extraction and classification. Only the joint effort by different teams, each one specialized in its own technique, could actually make it. In that sense, the contest was a success.

This classifier, which provides the best results ever obtained on this data set, can be considered in itself as a technical contribution of the contest.

VIII. CONCLUSION AND PERSPECTIVES

The contest provided some interesting conclusions and perspectives. They are summarized in the following items.

1) *Supervised versus unsupervised methods*. It was very interesting to see that the first uploaded results had been obtained with unsupervised methods. The results were fairly good (around 75%) but were outperformed by the supervised methods when they appeared a few weeks later. However, seeing these methods providing very fast and fairly good results was quite interesting.
2) *Dimension reduction*. Most of the proposed methods used a dimension reduction as a preprocessing. Most of them used the PCA, retaining various numbers of components. However, this step, with PCA or other methods, seems to be a must-do.
3) *Spatial and spectral features*. Several algorithms used both kinds of features. While the spectral information is easily extracted from the original spectra (directly or after some sort of dimension reduction), the spatial information remains a more tricky issue. Texture analysis and mathematical morphology provide some answers. Other ways to extract such a meaningful information are currently investigated. Similarly, mixing the spectral and the spatial information in the best possible way is also a clear direction for future research works.
4) *SVMs*. Almost all the best methods used some SVM-based classifiers. SVM really appeared as extremely suited for hyperspectral data, thus confirming the results presented in the recent abundant literature.
5) *NNs*. We must conclude by emphasizing that, similar to the 2007 contest, NNs have provided the best individual performances.

The final comment is on *decision fusion*. It was a great surprise and a very interesting point when we noticed that many submitted results had been obtained using different algorithms, meaning that the participants already performed a decision fusion before uploading their classification maps. This fusion "to the power of two" was also a clear sign that decision fusion is indeed a way to go for future research.

Of course, a crucial issue is the algorithm used for the fusion. The simplest solution consists in performing a majority vote. Some participants used it; it was also used for the final result of the contest. However, this is clearly suboptimal. More advanced strategies require the definition of a reliability criterion [27], [28]. The solution used by Jordi Inglada and Emmanuel Christophe in the frame of the contest is both very smart and very inspiring: Using the confusion matrices automatically provided by the system may sound like a diversion of the contest. However, it is, as a matter of fact, absolutely reasonable for operational applications. Combining several classification results based on their performances on small areas, where a ground truth is available, corresponds to real application situations. In crisis situation, classification is usually performed by hand. Using such a system enables one to limit the human intervention only to a small portion of the image while keeping similar performances.

As a conclusion, the actual classification performances obtained at the end of the contest should not be considered as absolute values. The results were obtained after a few months of intense activity by all the participants and were obtained with one single data set. The accurate and reliable classification of hyperspectral images still needs some methodological developments. However, the conclusions, as discussed in this session, clearly point some ways for future research. Among them, decision fusion has doubtlessly demonstrated its outstanding ability.

REFERENCES

[1] L. Alparone, L. Wald, J. Chanussot, C. Thomas, P. Gamba, and L. M. Bruce, "Comparison of pansharpening algorithms: Outcome of the 2006 GRS-S data-fusion contest," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 10, pp. 3012–3021, Oct. 2007.
[2] F. Pacifici, F. Del Frate, W. J. Emery, P. Gamba, and J. Chanussot, "Urban mapping using coarse SAR and optical data: Outcome of the 2007 GRSS data fusion contest," *IEEE Geosci. Remote Sens. Lett.*, vol. 5, no. 3, pp. 331–335, Jul. 2008.
[3] V. W. Samawi and O. A. AL Basheer, "The effect of features reduction on different texture classifiers," in *Proc. 3rd ICIEA*, Jun. 2008, pp. 67–72.
[4] S. Kumar, J. Ghosh, and M. M. Crawford, "Best-bases feature extraction algorithms for classification of hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 7, pp. 1368–1379, Jul. 2001.
[5] A. C. Jensen and A. S. Solberg, "Fast hyperspectral feature reduction using piecewise constant function approximations," *IEEE Geosci. Remote Sens. Lett.*, vol. 4, no. 4, pp. 547–551, Oct. 2007.
[6] S. B. Serpico, M. D'Inca, F. Melgani, and G. Moser, "Comparison of feature reduction techniques for classification of hyperspectral remote sensing data," in *Proc. SPIE Image Signal Process. Remote Sens. VIII*, Jun. 2003, vol. 4885, pp. 347–358.
[7] T. K. Ho, J. J. Hull, and S. N. Srihari, "Decision combination in multiple classifier systems," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 16, no. 1, pp. 66–75, Jan. 1994.
[8] L. Lam and S. Y. Suen, "Application of majority voting to pattern recognition: An analysis of its behavior and performance," *IEEE Trans.*

Syst., Man, Cybern. A, Syst., Humans, vol. 27, no. 5, pp. 553–568, Sep. 1997.
[9] P. Hong, L. Chengde, L. Linkai, and Z. Qifeng, "Accuracy of classifier combining based on majority voting," in *Proc. IEEE ICCA*, May 30–Jun. 1, 2007, pp. 2654–2658.
[10] C. Bishop, *Neural Networks for Pattern Recognition*. London, U.K.: Oxford Univ. Press, 1995.
[11] F. Del Frate, F. Pacifici, G. Schiavon, and C. Solimini, "Use of neural networks for automatic classification from high-resolution images," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 4, pp. 800–809, Apr. 2007.
[12] M. Chini, F. Pacifici, W. J. Emery, N. Pierdicca, and F. Del Frate, "Comparing statistical and neural network methods applied to very high resolution satellite images showing changes in man-made structures at rocky flats," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 6, pp. 1812–1821, Jun. 2008.
[13] J. Serra, *Image Analysis and Mathematical Morphology*. New York: Academic, 1982.
[14] P. Soille, *Morphological Image Analysis*. Berlin, Germany: Springer-Verlag, 2004.
[15] M. Pesaresi and J. A. Benediktsson, "A new approach for the morphological segmentation of high-resolution satellite images," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 2, pp. 309–320, Feb. 2001.
[16] 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.
[17] J. A. Palmason, J. A. Benediktsson, and K. Arnason, "Morphological transformations and feature extraction for urban data with high spectral and spatial resolution," in *Proc. IGARSS*, 2003, pp. 470–472.
[18] J. A. Benediktsson, J. A. Palmason, and J. R. Sveinsson, "Classification of hyperspectral data from urban areas based on extended morphological profiles," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 3, pp. 480–490, Mar. 2005.
[19] A. Plaza, P. Martinez, J. Plaza, and R. Perez, "Dimensionality reduction and classification of hyperspectral image data using sequences of extended morphological transformations," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 3, pp. 466–479, Mar. 2005.
[20] S. R. Sternberg, "Grayscale morphology," *Comput. Vis. Graph. Image Process.*, vol. 35, no. 3, pp. 333–355, Sep. 1986.
[21] R. Collobert, S. Bengio, and J. Mariéthoz, "Torch: A modular machine learning software library," IDIAP, Martigny, Switzerland, Tech. Rep. RR 02-46, 2002.
[22] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*. New York: Wiley-Interscience, 2000.
[23] S. Prasad and L.-M. Bruce, "Decision fusion with confidence-based weight assignment for hyperspectral target recognition," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 5, pp. 1448–1456, May 2008.
[24] P. M. Mather, *Computer Processing of Remotely-Sensed Images*. New York: Wiley, 1999.
[25] J. C. Platt, "Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods," in *Advances in Large Margin Classifiers*. Cambridge, MA: MIT Press, 1999.
[26] C. Thiel, F. Giacco, F. Schwenker, and G. Palm, "Comparison of neural classification algorithms applied to land cover mapping," in *Proc. 18th Italian WIRN*, 2008, pp. 254–263.
[27] J. Chanussot, G. Mauris, and P. Lambert, "Fuzzy fusion techniques for linear features detection in multi-temporal SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 3, pp. 1292–1305, May 1999.
[28] 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, pp. 2828–2838, Oct. 2006.

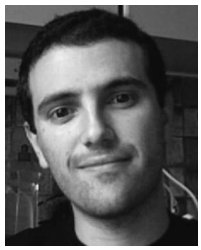
Giorgio Licciardi received the M.S. degree in telecommunications engineering from Tor Vergata University, Rome, Italy, in 2005, defending a thesis on the use of neural networks for image information mining from satellite imagery. He is currently working toward the Ph.D. degree in geoinformation at Tor Vergata University.

He is a Research Fellow on satellite data classification using neural networks algorithms. He is also a European Space Agency Category-1 Principal Investigator for Earth observation data. He is the coauthor of several scientific publications regarding the use of satellite data for the production of land-cover and change-detection maps, which is also the main subject of his Ph.D. degree.

Fabio Pacifici (S'03) received the Laurea (B.S.) (*cum laude*) and the Laurea Specialistica (M.S.) (*cum laude*) degrees in telecommunication engineering from Tor Vergata University, Rome, Italy, in 2003 and 2006, respectively. He is currently working toward the Ph.D. degree in geoinformation in the Earth Observation Laboratory, Tor Vergata University, and collaborates with the University of Colorado, Boulder.

His research activities include remote sensing image processing, analysis of multitemporal data, and data fusion. In particular, his interests are related to the development of novel classification and change-detection techniques for urban remote sensing applications using very high resolution optical and synthetic aperture radar imagery. He serves as a Referee for the *European Association for Signal Processing Journal on Image and Video Processing*.

Mr. Pacifici was the recipient of the 2009 Joint Urban Remote Sensing Event Student Paper Competition. He serves as a Referee for the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, IEEE GEOSCIENCE AND REMOTE SENSING LETTERS, and IEEE TRANSACTIONS ON IMAGE PROCESSING.



Devis Tuia (S'07) received the Diploma in geography from the University of Lausanne, Lausanne, Switzerland, in 2004 and the Master of Advanced Studies degree in environmental engineering from the Federal Institute of Technology of Lausanne, Lausanne, in 2005. He is currently working toward the Ph.D. degree at the Institute of Geomatics and Analysis of Risk, University of Lausanne, in the field of machine learning and its applications to urban remote sensing.

Mr. Tuia serves as a Referee for the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING and the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS.

Saurabh Prasad (S'03–M'09) received the Ph.D. degree in electrical engineering from the Mississippi State University, Starkville, in 2008.

He is currently an Assistant Research Professor with the Geosystems Research Institute and an Adjunct Professor with the Electrical and Computer Engineering Department, Mississippi State University. His research interest includes the use of information fusion techniques for designing robust statistical pattern classification algorithms for hyperspectral remote sensing systems.

Terrance West (S'06) received the B.S.E. and M.S. degrees in electrical and computer engineering from the Mississippi State University, Starkville, in 2004 and 2006, respectively, where he is currently working toward the Ph.D. degree in electrical and computer engineering.

He is currently a Graduate Research Assistant with the GeoResources Institute, Mississippi State University's High Performance Computation Collaboratory. His research interests include signal processing and statistical pattern recognition and their application to automated target recognition and land-cover classification in hyperspectral remote sensing applications. In particular, his current research work involves the use of wavelet-based feature extraction and selection techniques in a multiclassifier framework for designing statistical pattern classification algorithms for hyperspectral remote sensing systems.

Ferdinando Giacco received the Laurea degree in physics from the University of Napoli, Napoli, Italy, in 2005 and the Ph.D. degree from the University of Salerno, Salerno, Italy, in 2009.

His research interests are complex systems and statistical mechanics.

Christian Thiel received the M.Sc. degree (German Diplom) from University of Ulm, Germany, in 2004.

He is currently with the Institute of Neural Information Processing, University of Ulm, Ulm, Germany, advancing several research projects and his dissertation, which is to be completed at the end of 2009. His research interest includes pattern recognition, particularly multiple classifier systems and classification under uncertainty.

Jordi Inglada received the Telecommunications Engineer degree from both the Universitat Politècnica de Catalunya, Barcelona, Spain, and the École Nationale Supérieure des Télécommunications de Bretagne, Brest, France, in 1997 and the Ph.D. degree in signal processing and telecommunications in 2000 from Université de Rennes 1, Rennes, France.

He is currently with the Centre National d'Études Spatiales (French Space Agency), Toulouse, France, working in the field of remote sensing image processing. He is in charge in the development of image processing algorithms for the operational exploitation of Earth observation images, mainly in the fields of image registration, change detection, and object recognition.

Emmanuel Christophe received the Engineering degree in telecommunications from the École Nationale Supérieure des Télécommunications de Bretagne, Brest, France, and the D.E.A. degree (with honors) in telecommunications and image processing from the University of Rennes 1, Rennes, France, both in 2003 and the Ph.D. degree in hyperspectral image compression and image quality from the École Nationale Supérieure de l'Aéronautique et de l'Espace, Toulouse, France, and the University of Toulouse, Toulouse, in 2006.

He was a Visiting Scholar at Rensselaer Polytechnic Institute, Troy, NY, in 2006. From 2006 to 2008, he was a Research Engineer with the Centre National d'Etudes Spatiales (French Space Agency), Toulouse, focusing on information extraction for high-resolution optical images. Since that time, he has also been deeply involved in the development of the open-source Orfeo Toolbox. In October 2008, he joined the Centre for Remote Imaging, Sensing and Processing, National University of Singapore, Singapore, Singapore, where he is currently tackling new challenges for remote sensing in tropical areas. His research interests include image and video compression, as well as image processing and computer vision for remote sensing.

Jocelyn Chanussot (SM'04) received the M.Sc. degree in electrical engineering from the Grenoble Institute of Technology (INPG), Grenoble, France, in 1995 and the Ph.D. degree from the University of Savoie, Annecy, France, in 1998.

In 1999, he was with the Geography Imagery Perception Laboratory, Delegation Generale de l Armement (French National Defense Department). Since 1999, he has been with INPG, where he was an Assistant Professor from 1999 to 2005, an Associate Professor from 2005 to 2007, and is currently a Professor of signal and image processing. He is currently conducting his research at the Laboratoire Grenoblois de l'Image, de la Parole, du Signal et de l'Automatique. His research interests include image analysis, multicomponent and hyperspectral image processing, nonlinear filtering, and data fusion in remote sensing.

Dr. Chanussot is an Associate Editor of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING. He was an Associate Editor of the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS (in 2005–2007) and *Pattern Recognition* (in 2006–2008). He is the Chair (in 2009–2011) and the Cochair (in 2005–2008) of the IEEE Geoscience and Remote Sensing Society (GRS-S) Data Fusion Technical Committee and a member of the Machine Learning for Signal Processing Technical Committee of the IEEE Signal Processing Society (in 2006–2008). He is the founding President of the IEEE GRS-S French chapter (in 2007) and a member of the IEEE GRS-S Administrative Committee (in 2009–2011). He is the General Chair of the first IEEE GRS-S Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing and the Program Cochair of the 2009 IEEE International Workshop on Machine Learning for Signal Processing.

Paolo Gamba (S'91–M'93–SM'00) received the Laurea (*cum laude*) and Ph.D. degrees in electronic engineering from the University of Pavia, Pavia, Italy, in 1989 and 1993, respectively.

He is currently an Associate Professor of telecommunications with the University of Pavia. He published more than 60 papers on international peer-reviewed journals and presented more than 150 papers in workshops and conferences. He is or has been the Guest Editor of special issues of the International Society for Photogrammetry and Remote Sensing (ISPRS) *Journal of Photogrammetry and Remote Sensing* and the *International Journal of Information Fusion and Pattern Recognition Letters* on the topics of urban remote sensing, remote sensing for disaster management, and pattern recognition in remote sensing applications.

Dr. Gamba is or has been the Guest Editor of special issues of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING. He has been Chair of Technical Committee 7 "Pattern Recognition in Remote Sensing" of the International Association for Pattern Recognition (IAPR) from October 2002 to October 2004 and Chair of the Data Fusion Committee of the IEEE Geoscience and Remote Sensing Society from October 2005 to May 2009. Since January 2009, he serves as Editor-in-Chief of the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS. He will be the Technical Cochair of the 2010 IEEE Geoscience and Remote Sensing Symposium scheduled for July 2010 in Honolulu, HI. He is the Organizer and the Technical Chair of the biennial Geoscience and Remote Sensing Society/ISPRS Joint Workshops on "Remote Sensing and Data Fusion over Urban Areas" from 2001 to 2009. The last event he chaired was the 2009 Joint Urban Remote Sensing Symposium, held in Shanghai, China in May 2009.