Spectral and Spatial Classification of High Resolution Urban Satellites Images using Haralick features and SVM with SAM and EMD distance Metrics

Aissam Bekkari, Soufian Idbraim, Azeddine Elhassouny, Driss Mammass Mostapha El Yassa

IRF – SIC laboratory,
Faculty of sciences
Agadir -Morocco

Danielle Ducrot

Cesbio, Toulouse- France

ABSTRACT

The classification of remotely sensed images knows a large progress taking into consideration the availability of images with different resolutions as well as the abundance of classification's algorithms. Support Vector Machines (SVMs) are a group of supervised classification algorithms that have been recently used in the remote sensing field, a number of works have shown promising results by the fusion of spatial and spectral information using SVM.

For this purpose, we propose a methodology allowing to combine these two information. The SVM classification was conducted using a combination of multi-spectral features and Haralick texture features as data source. We have used homogeneity, contrast, correlation, entropy and local homogeneity, which were the best texture features to improve the classification algorithm. Two kernels have been considered, the RBF kernel based on Euclidean minimum distance (EMD) and a RBF kernel based on the Spectral Angle Mapper (SAM).

The proposed approach was tested on common scenes of urban imagery. Results showed, especially with the use of Haralick texture features, that SVMs using RBF kernel based on EMD outperform the SVMs with RBF kernel based on SAM in term of the global accuracy and Kappa coefficeint. The experimental results indicate a mean accuracy value of 93.406% for EMD kernel and 92.896% for SAM kernel which is very promising.

Keywords

SVM, Classification, SAM, EMD, Haralick features, Satellite image, Spectral and spatial information, GLCM.

1. INTRODUCTION

With the commercial emergence of the optical satellite images of sub-metric resolution (Ikonos, Quickbird) the realization as well as the regular update of numerical maps with large scales becomes accessible and increasingly frequent. The classification of such images follows the same principle for other image types, and it is a method of analysis of data that aims to separate the image into several classes in order to gather the data in homogeneous subsets, which show common characteristics. It aims to assign to each pixel of the image a label which represents a theme in the real study area (e.g. vegetation, water, built, etc) [1].

Several classification algorithms have been developed since the first satellite image was acquired in 1972 [2-4]. Among the most popular and widely used is the maximum likelihood classifier [5]. It is a parametric approach that assumes the class signature in normal distribution. Although this assumption is generally valid, it is invalid for classes consisting of several subclasses or classes having different spectral features [6]. To overcome this problem, some non-parametric classification techniques such as artificial neural networks, decision trees and Support vector machines (SVM) have been recently introduced.

SVM is a group of advanced machine learning algorithms that have seen increased use in land cover studies [7, 8]. One of the theoretical advantages of the SVM over other algorithms (decision trees and neural networks) is that it is designed to search for an optimal solution to a classification problem whereas decision trees and neural networks are designed to find a solution, which may or may not be optimal. This theoretical advantage has been demonstrated in a number of studies where SVM generally produced more accurate results than decision trees and neural networks [5, 9]. SVMs have been used recently to map urban areas at different scales with different remotely sensed data. High or medium spatial resolution images (e.g., IKONOS, Quickbird, Landsat (TM)/ (ETM+), SPOT) have been widely employed on urban land use classification for individual cities in building extraction, road extraction and other man-made objects extraction [10, 11].

Moreover, in hyperspectral image analysis, especially in application for classifying and detecting, spectral characterization plays a more and more crucial role. To determine spectral variability, similarity, discrimination, and divergence, many spectral measure criteria which calculate different distance metrics have been proposed over the past few decades, including spectral angle mapper (SAM) [12], spectral correlation mapper (SCM) [13], spectral information measure (SIM) [14], Euclidean minimum distance (EMD) [15], spectral gradient angle (SGA) [16], and band add-on spectral angle mapper (BAO-SAM) [17]. Appropriate distance metrics employed in hyperspectral data processing for classification and detection application can obtain the best result through describing spectral characteristics in mathematical or physical meaning properly.

SAM, as one of the most important and widely used spectral distance metrics in hyperspectral data processing for classification and material identification, is implemented and applied in many research and commerce fields, however, as a physically-based spectral distance metrics [18].

On the other hand, the consideration of the spatial aspect in the spectral classification remains very important, for this case, Haralick described methods for measuring texture in gray-scale images, and statistics for quantifying those textures [19]. It is the hypothesis of this research that Haralick's Texture Features and statistics as defined for gray-scale images can be modified to incorporate spectral information, and that these Spectral Texture Features will provide useful information about the image. It is shown that texture features can be used to classify general classes of materials, and that Spectral Texture Features in particular provide a clearer classification of land cover types than purely spectral methods alone.

In this paper, multiclass SVMs are investigated for the classification of multispectral high resolution satellite images. The proposed method consists in combining spatial and spectral information to obtain a better classification using Haralick Features [19].

Two kernels are compared: The first one is based on Euclidean Minimum Distance. It is the RBF with L2-norm distance. The second one is based on Spectral Angle Mapper. It basically computes the angle between two vectors in the vector space.

This paper is organized as follows. In the second section, we discuss the extraction of spatial and spectral information especially the Grey-Level Co-occurrence Matrix (GLCM) and Haralick texture features used in experiments. In Section 3, we give outlines on the used classifier: Support Vector Machines (SVM). In Section 4, the two distance metrics used in experiments are outlined, and Data. Experimental scheme and experimental results are discussed in Section 5. Finally, conclusions and perspectives are given in Section 6.

2. EXTRACTION OF INFORMATION

2.1 Spectral information

The most used classification methods for the multispectral data consider especially the spectral dimension. The set of spectral values of each pixel is treated as a vector of attributes which will be directly employed as an entry of the classifier. According to Fauvel [20] this allows a good classification based on the spectral signature of each area. However, this does not take into account the spatial information represented by the various structures in the image.

2.2 Spatial information

Information in a remote sensed image can be deduced basing on their textures. A human analyst is able to distinguish man-made features from natural features in an image based on the 'regularity' of the data. Straight lines and regular repetitions of features hint at man-made objects. This spatial information is useful in distinguishing the different field in the remote sensed image.

Many approaches were developed for texture analysis. According to the processing algorithms, three major categories, namely, structural, spectral, and statistical methods are common ways for texture analysis.

Haralick [21] suggested the use of GLCM to extract second order statistics from an image. GLCMs have been used very successfully for texture classification in evaluations [22].

It's one of the most widely used methods [23], which is a powerful technique for measuring texture features; it contains the relative frequencies of the two neighbouring pixels separated by a distance on the image (Fig 1).

The size of the co-occurrence matrix equals to the number of the image gray levels, also the dynamics of the image is usually small (typically, 8 gray levels) in order not to work with too large matrices.

The distribution in the matrix will depend on the angular and distance relationship between pixels. Varying the vector used allows the capturing of different texture characteristics. Once the GLCM has been created, various features can be computed from it. These have been classified into four groups: visual texture characteristics, statistics, information theory and information measures of correlation [21, 24].

0	0	1	1
0	0	1	1
0	2	2	2
2	2	3	3

$$\begin{pmatrix} 4 & 2 & 1 & 0 \\ 2 & 4 & 0 & 0 \\ 1 & 0 & 6 & 1 \\ 0 & 0 & 1 & 2 \end{pmatrix} \qquad \begin{pmatrix} 2 & 0 & 3 & 0 \\ 0 & 0 & 2 & 2 \\ 3 & 2 & 0 & 0 \\ 0 & 2 & 0 & 0 \end{pmatrix}$$

(a) (b) (c)

Fig 1. (a) Image 4 × 4 with 4 gray levels, (b) co-occurrence matrix for a displacement of d = (1, 0) and (c) co-occurrence matrix for a displacement of d = (0, 2)

Even small, a co-occurrence matrix represents a substantial amount of data that is not easy to handle. This is why Haralick uses these matrices to develop a number of spatial indices that are easier to interpret.

Haralick assumed that the texture information is contained in the co-occurrence matrix, and texture features are calculated from it. A large number of textural features have been proposed starting with the original fourteen features (f_1 to f_{14}) described by Haralick et al [25], however only some of these features are in wide use. Wezska [26] used four of Haralick features (f_1, f_2, f_5, f_8). Conners and Harlow [27] use five features (f_1, f_2, f_3, f_4, f_5). Conners, Trivedi and Harlow [28] introduced two new features which address a deficiency in the Conners and Harlow set ($f_1, f_2, f_4, f_5, f_6, f_7$)

We found that the five features used by Conners and Harlow are commonly used taking into account that the fourteen are much correlated with each other, and that the five sufficed to give good results in classification [29].

In this work, we have used these five features: homogeneity (E), contrast (C), correlation (Cor), entropy (H) and local homogeneity (LH), and co-occurrence matrix are calculated for four directions: 0°, 45°, 90°and 135° degrees.

Let us recall their definitions:

$$E = \sum_i \sum_j (M(i,j))^2 \tag{1}$$

$$C = \sum_{k=0}^{m-1} k^2 \sum_{|i-j|=k} M(i,j) \tag{2}$$

$$Cor = \frac{1}{\sigma_i \sigma_j} \sum_i \sum_j (i-\mu_i)(j-\mu_j) M(i,j) \tag{3}$$

Where ***M(i, j)*** is normalized co-occurrence matrix, μ_i and σ_i are the horizontal mean and the variance, and μ_j and σ_j are the vertical statistics.

$$H = \sum_i \sum_j M(i,j) \log(M(i,j)) \tag{4}$$

$$LH = \sum_i \sum_j \frac{M(i,j)}{1+(i-j)^2} \tag{5}$$

Each texture measure can create a new band that can be incorporated with spectral features for SVM classification.

3. SVM CLASSIFICATION

In this section, we briefly describe the general mathematical formulation of SVMs introduced by Vapnik [30, 31]. Starting from the linearly separable case, optimal hyperplanes are introduced. Then, the classification problem is modified to handle non-linearly separable data and a brief description of multiclass strategies is given.

3.1 Linear SVM

For a two-class problem in a *n*-dimensional space R^n, we assume that *l* training samples $x_i \in R^n$ are available with their corresponding labels *yi* = ±1, S = {(*xi*, *yi*) | i ∈ [1, *l*]}. The SVM method consists of finding the hyperplane that maximizes the margin, i.e., the distance to the closest training data points for both classes [32]. Noting $w \in R^n$ as the normal vector of the hyperplane and $b \in R$ as the bias, the hyperplane H_p is defined as:

$$\langle w, x \rangle + b = 0, \forall x \in H_p \tag{6}$$

Where $\langle w, x \rangle$ is the inner product between *w* and *x*. If $x \notin Hp$ then $f(x) = \langle w, x \rangle + b$ is the distance of *x* to H*p*. The sign of *f* corresponds to decision function *y = sgn (f(x))*.

Finally, the optimal hyperplane has to maximize the margin: $2/\|w\|$. This is equivalent to minimize $\|w\|/2$ and leads to the following quadratic optimization problem:

$$\min\left[\frac{\|w\|^2}{2}\right] \quad \text{subject to } y_i(\langle w, x_i \rangle + b) \geq 1 \ \ \forall i \in [1,l] \tag{7}$$

For non-linearly separable data, the optimal parameters (*w*, *b*) are found by solving:

$$\min\left[\frac{\|w\|^2}{2} + C\sum_{i=1}^{l} \xi_i\right] \quad \text{subject to } y_i(\langle w, x_i \rangle + b) \geq 1 - \xi_i, \xi_i \geq 0 \ \ \forall i \in [1,l] \tag{8}$$

Where the constant *C* control the amount of penalty and ξ_i are *slack* variables which are introduced to deal with misclassified samples (Fig 2). This optimization task can be solved through its Lagrangian dual problem:

$$\max_{\alpha} \sum_{i=1}^{l} \alpha_i - \frac{1}{2}\sum_{i,j=1}^{l} \alpha_i \alpha_j y_i y_j \langle x_i, x_j \rangle \quad \text{subject to } 0 \leq \alpha_i \leq C \quad \forall i \in [1,l] \quad \sum_{i=1}^{l} \alpha_i y_i = 0 \tag{9}$$

Finally:

$$w = \sum_{i=1}^{l} \alpha_i y_i x_i \tag{10}$$

The solution vector is a linear combination of some samples of the training set, whose α_i is non-zero, called Support Vectors. The hyperplane decision function can thus be written as:

$$y_u = \text{sgn}\left(\sum_{i=1}^{l} y_i \alpha_i \langle x_u, x_i \rangle + b\right) \tag{11}$$

Where x_u is an unseen sample.

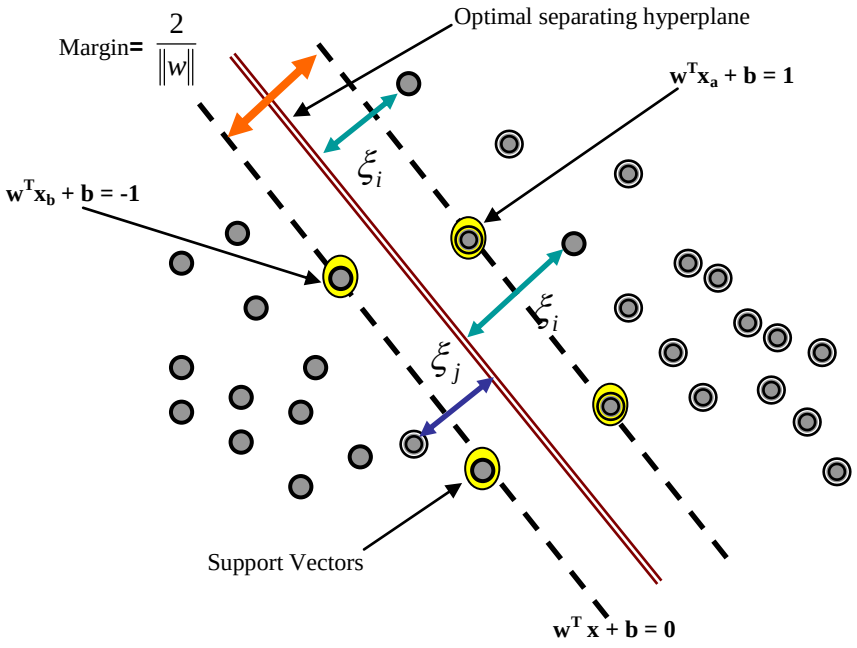


Fig 2. Classification of a non-linearly separable case by SVMs.

3.2 Non-Linear SVM

Using the Kernel Method, we can generalize SVMs to non-linear decision functions. With this technique, the classification capability is improved. The idea is as follows. Via a non-linear mapping Φ, data are mapped onto a higher dimensional space F (Fig 3):

$$\Phi : R^n \rightarrow F \qquad x \mapsto \Phi(x) \tag{12}$$

The SVM algorithm can now be simply considered with the following training samples: $\Phi(S) = \{(\Phi(x_i), y_i) \mid i \in [1, l]\}$. It leads to a new version of the hyperplane decision function where the scalar product is now: $\langle \Phi(x_i), \Phi(x_j) \rangle$. Hopefully, for some kernels function k, the extra computational cost is reduced to:

$$\langle \Phi(x_i), \Phi(x_j) \rangle = k(x_i, x_j) \tag{13}$$

The kernel function k should fulfill Mercers' conditions.

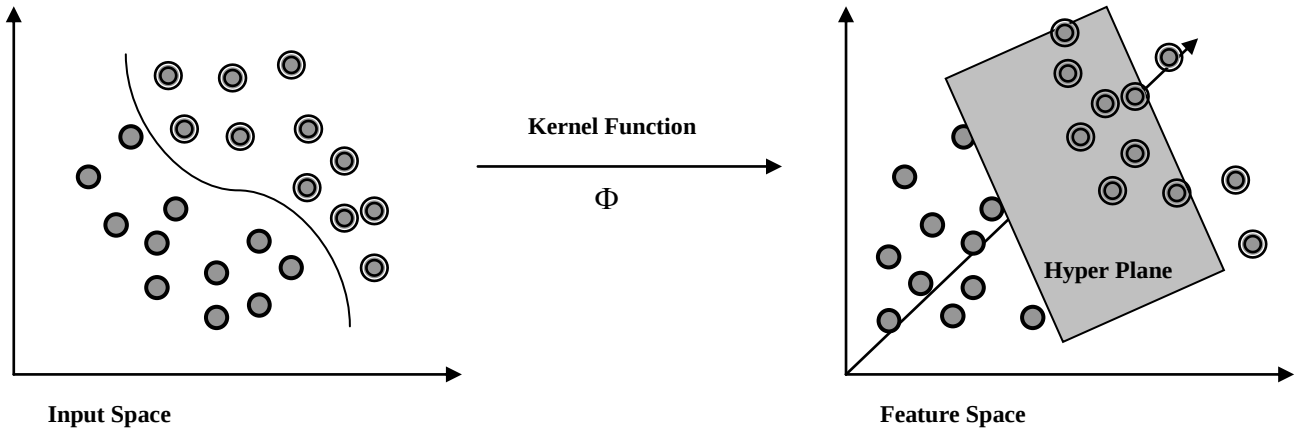


Fig 3. Mapping the Input Space into a High Dimensional Feature Space with a kernel function

With the use of kernels, it is possible to work implicitly in F while all the computations are done in the input space. The classical kernels used in remote sensing are the polynomial kernel and the Gaussian radial basis function:

$$k_{poly}(x_i, x_j) = \left[(x_i \cdot x_j) + 1\right]^p \tag{14}$$

$$k_{gauss}(x_i, x_j) = \exp\left[-\gamma \left\| x_i - x_j \right\|^2\right] \tag{15}$$

3.3 Multiclass SVMs

SVMs are designed to solve binary problems where the class labels can only take two values: ±1. For a remote sensing application, several classes are usually of interest. Various approaches have been proposed to address this problem [33]. They usually combine a set of binary classifiers. Two main approaches were originally proposed for a k-classes problem.

- ***One versus the Rest***: k binary classifiers are applied on each class against the others. Each sample is assigned to the class with the maximum output.
- ***Pairwise Classification***: $k(k-1)/2$ binary classifiers are applied on each pair of classes. Each sample is assigned to the class getting the highest number of votes. A vote for a given class is defined as a classifier assigning the pattern to that class.

4. DISTANCE METRICS

In essence, a distance metric is a mathematical operator that conveys how similar two members of a set are with a single scalar value, based on a notion of similarity [17, 34].

Different metrics employ alternative notions of similarity; consequently, each metric uniquely translates the phenomenology observed by a sensor into a scalar. The notion of similarity shared by the spectra can be measured differently, depending on the metric that is used to compare them.

In this section, we discuss the two most prominent distance metrics in hyperspectral processing: the Euclidean Minimum Distance and the Spectral Angle Mapper. Each metric provides a unique measure of distance from two complementary viewpoints of geometry (Fig 4).

4.1 Euclidean minimum distance

The Fig.4 (a) shows that EMD measures the shortest distance between two vectors x and y, and is defined as:

$$\Delta(x, y) = \|x - y\| \tag{16}$$

From the definition of EMD in (16), EMD possesses properties that make it distinct from SAM, which are: *Invariance to Unitary Coordinate Transformation*, *Additivity*, and *Monotonicity*.

So the contrast between two signals measured by EMD increases with the number of bands. Furthermore, the additivity of EMD confirms that the amount of contrast between two signals increases with additional bands independently to other bands value. In short, the greatest contrast between two spectra is necessarily achieved with EMD by using every available band.

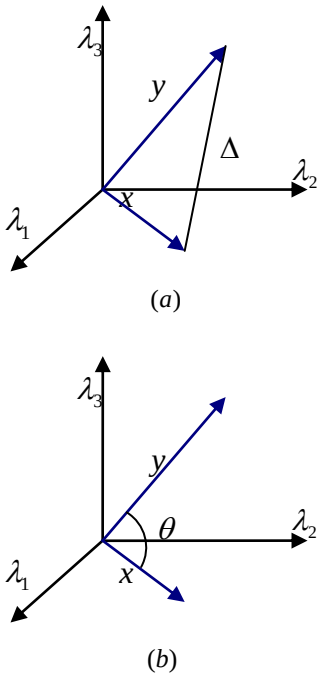


Fig 4. (a) and (b) represent respectively the Distance and the Spectral angle between two spectra, *y* = target spectrum, *x* = reference spectrum, using three bands λ_1, λ_2, λ_3.

4.2 Spectral angle mapper

For each selected pixel in a hyperspectral image, namely unknown spectrum x, and a reference spectrum y chosen from spectral libraries, SAM algorithm qualifies the spectral angle θ to determine the similarity between them by applying the following equation:

$$\theta(x, y) = \arccos\left(\frac{\langle x, y\rangle}{\|x\|\|y\|}\right) \quad 0 \le \theta \le \frac{\pi}{2} \tag{17}$$

Where $\langle .,. \rangle$ is the dot product operator, and $\|.\|$ is the 2-norm.

A pair of three-dimensional spectra and the corresponding spectral angle calculated by SAM algorithm are shown in Fig.4 (b) In the SAM algorithm, smaller spectral angle indicates more similar to the reference spectrum.

From its mathematical definition in (17), SAM possesses unique properties that distinguish it from EMD. These are *Invariance to Multiplicative Scaling*, *Nonadditivity* and *Nonmonotonicity*.

So the addition of more spectral bands does not always guarantee an increase in angle.

5. EXPERIMENTAL RESULTS

5.1 Experiments

The idea for a good spectral classification of the pixels is to directly consider the value of pixels image as input data of the

classifier, and each texture measure will create a new band that will be incorporated with this spectral information to use jointly spatial and spectral information.

The proposed workflow has two main tasks, we start with the extraction of spectral and spatial information, so we compute Grey Level Co-occurrence Matrix (GLCM) to extract Haralick texture features that we add to spectral information, and then the result will be used as an input to SVM classifier (Fig 5).

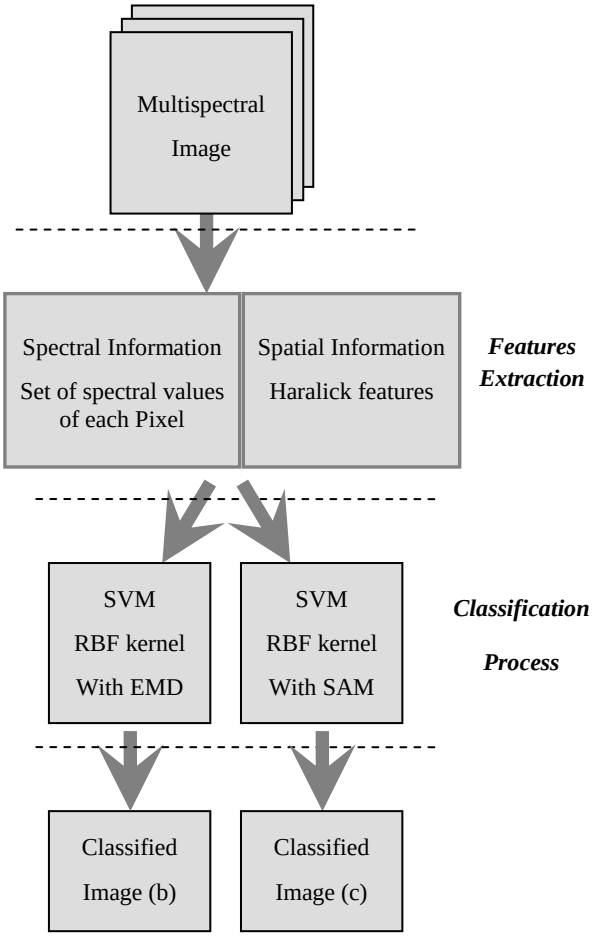


Fig 5. A representative illustration of the workflow

The performance of SVM varies depending on the choice of the kernel function and its parameters. For RBF kernel, two parameters, which are regularization parameter (C) and kernel width (γ), need to be defined. It is not clear which pairs of parameter produce the best classification result for a given data set. Therefore, optimum parameter search must be performed [35].

In this work, the parameters of RBF kernel were determined by a grid search method using cross validation approach. The main idea behind the grid search method is that different pairs of parameters are tested and the one with the highest cross validation accuracy is selected. The method is conducted in two steps. In the first step, a coarser grid is applied with an exponentially growing sequence of (C, γ). In the second step, after identifying the optimal region on the grid, the finer grid search is executed. The results are used to perform the final training process [35, 36].

For classification, we have used SVMlight wich is an implementation of Support Vector Machines (SVMs) in C language [37]. We have modified the source code so as to adapt it with the used kernels.

RBF can be written as follows [38]: $K(\mathbf{x}, y) = f(d(x, y))$ where d is a metric on R^n and f is a function on R^+_0 . For the Gaussian RBF, $f(t) = exp(-\gamma t^2)$, $t \in R^+_0$, and $d(x, y)=||x-y||$, i.e., the Euclidean minimum distance. As mentioned in [17], Euclidean distance is not scale invariant, however due to atmospheric attenuation or variation in illumination, spectral energy can be different for two samples even if they belong to the same class. To handle such a problematic case, scale invariant metrics can be considered. *Spectral Angle Mapper* (SAM) is a well known scale invariant metric, it has been widely used in many remote sensing problems and it has been shown to be robust to variations in spectral energy [17]. This metric θ focuses on the angle between two vectors (17).

In this paper, we compare RBF kernels with the Euclidean Distance (15) and kernel based on Spectral Angle Mapper (18).

$$K_{SAM}(x, y) = \exp\left[-\gamma\theta(x, y)^2\right] \quad (18)$$

Both kernels fulfill Mercer conditions and optimal hyperplanes can therefore be found.

5.2 Data

The first image used in classification is a sample of high resolution Quickbird satellite image. Its size is 240x360 pixels. It represents scene urban areas. We dispose of four spectral bands: blue, green, red and near infrared. We can see in Fig.6 (a) a representation of this image.

The second test image is another sample of Quickbird satellite image with exactly the same properties except the size, 500x280 pixels. The scene does contain also urban areas. The original image is represented in Fig.7 (a).

In order to complete this collection the last image is a sample of high resolution Ikonos satellite image. It has also four spectral bands: red, blue, green and near infrared, its size is 600x800 pixels. This image is represented in Fig.8 (a).

Table 1. Different classes

Class N°	***Class name***	**Train samples**		
		Image 1	**Image 2**	**Image 3**
1	Asphalt	1 592	753	1 386
2	Green area	2 252	1 680	480
3	Tree	880	519	196
4	Soil	176	1 387	813
5	Building	4 217	1 282	920
6	Shadow	1 280	808	336
Total		10 397	6 429	4 131

We will have two files for each image, "TrainFile.dat" and "TestFile.dat" respectively for learning and for classification, divided on six classes as described in Table 1.

5.3 The results

To the file that contains spectral and spatial information obtained from the original image ((a) in Fig 6, 7 and 8), we apply an SVM classification with RBF kernel based on EMD (15) then we apply, on the same file, an SVM classification with RBF kernel based on SAM (18).

Thematic maps provided by SVMs classification using spectral information and Haralick features with the EMD and SAM kernel are shown in Fig.6, 7 and 8, (b) and (c), respectively.

The fusion of spectral information and Haralick features with SAM metric gives us the classification maps (c) respectively in Fig 6, 7 and 8.

The results have progressed with the combined use of spectral and spatial information. In addition, a visual analysis of the classification maps shows those areas more homogeneous for the map obtained with the proposed SVM using spectral and Haralick features with both SAM and EMD metrics.

In conclusion both of those methods matches well with an urban land cover map in terms of smoothness of the classes; and it also represents more connected classes mainly when using SVM classification with SAM, it gives us approximately the same result when using SVM classification with EMD in terms of global accuracy.

Table 2 summarizes the results obtained using the EMD and the SAM RBF kernels. These values were extracted from the confusion matrix, table 3 and table 4 present examples of the confusion matrix for the first used image respectively for EMD and SAM distance Metrics. The overall accuracy is the percentage of correctly classified pixels. Kappa coefficient is another criterion classically used in remote sensing classification to measure the degree of agreement and takes into account the correct classification that may have been obtained "by chance" by weighting the measured accuracies.

The use of the SAM kernel gives slightly degraded classification results for the overall accuracy and the Kappa coefficient. However, with all of the accuracies over 90%, this kernel seems also promising for the classification of remotely sensed images.

Table 2. Classification Accuracies for the EMD and the SAM RBF kernel.

Test images	Overall accuracy		Kappa Coefficient	
	EMD	SAM	EMD	SAM
Image 1 Fig.6 (a)	94.13%	93. 58%	0.93	0.92
Image 2 Fig.7 (a)	93.96%	93.29%	0.92	0.90
Image 3 Fig.8 (a)	92.13%	91.82%	0.91	0.89

6. CONCLUSION

Addressing the classification of high resolution satellite images from urban areas, we have presented two RBF kernels taking simultaneously the spectral and the spatial information into account the spectral values (Haralick features).

Two kernels have been compared, the well known EMD RBF kernel and a kernel based on the spectral angle mapper. From our experiments, both gave excellent results in terms of classification accuracy, the EMD kernel slightly outperforms the SAM kernel; however it remains to improve even more these results.

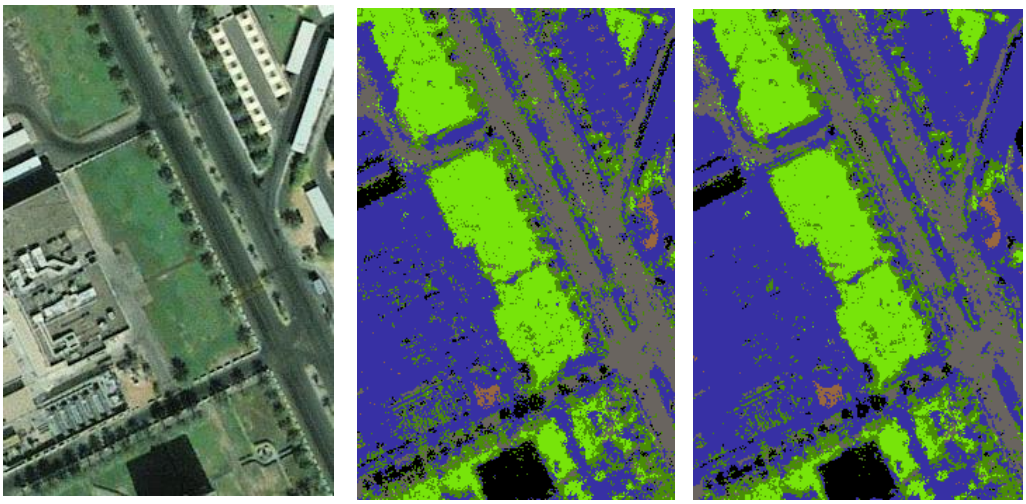
As a perspective of this work, we will be concentrating on the study of the kernel choice in order to determine the appropriate one, for this type of image classification. Another aspect of the proposed method which should be improved is the set of features used for the classification. We think that it is not possible, from a spectral and spatial SVM classification point of view, to eliminate a significant number of features among those used in our method; some feature selection methods should be tested.

Table 3. Confusion matrix results (%) for SVM classification with EMD (image 1 Fig.6 (a)) Global accuracy = 94.13%

Class name	Asphalt	Green area	Tree	Soil	Building	Shadow
Asphalt	**96.02**	0.34	1.92	0	0.62	1.09
Green area	1.13	**96.97**	0	0	1.53	0.36
Tree	0.18	1.22	**89.57**	1.98	2.53	4.51
Soil	0	0.08	0.03	**96.37**	3.51	0
Building	2.04	1.16	4.91	0.08	**91.8**	0
Shadow	0.62	0.22	3.57	1.56	0	**94.03**

Table 4. Confusion matrix results (%) for SVM classification with SAM (image 1 Fig.6 (a))
Global accuracy = 93.58%

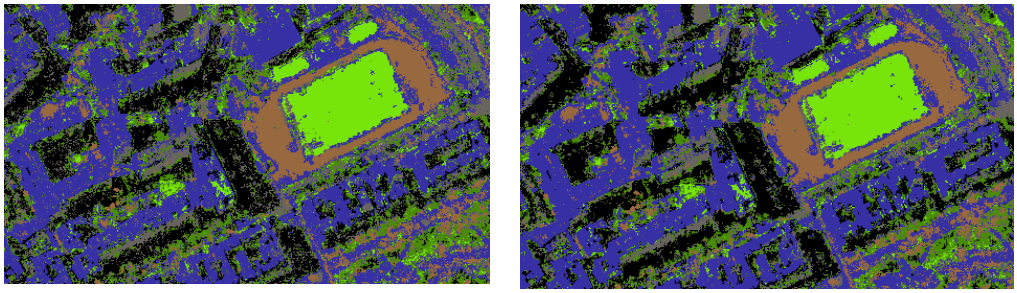
Class name	Asphalt	Green area	Tree	Soil	Building	Shadow
Asphalt	**95,02**	1,41	1,92	0	1,63	0,02
Green area	0,73	**95,39**	0	1,08	1,54	1,26
Tree	0,28	1,07	**88,82**	2,5	3,42	3,91
Soil	1,34	0,95	0	**95,87**	1,34	0,5
Building	2,01	0,76	4,69	0,47	**92,07**	0
Shadow	0,62	0,42	4,57	0,08	0	**94,31**



(a) (b) (c)

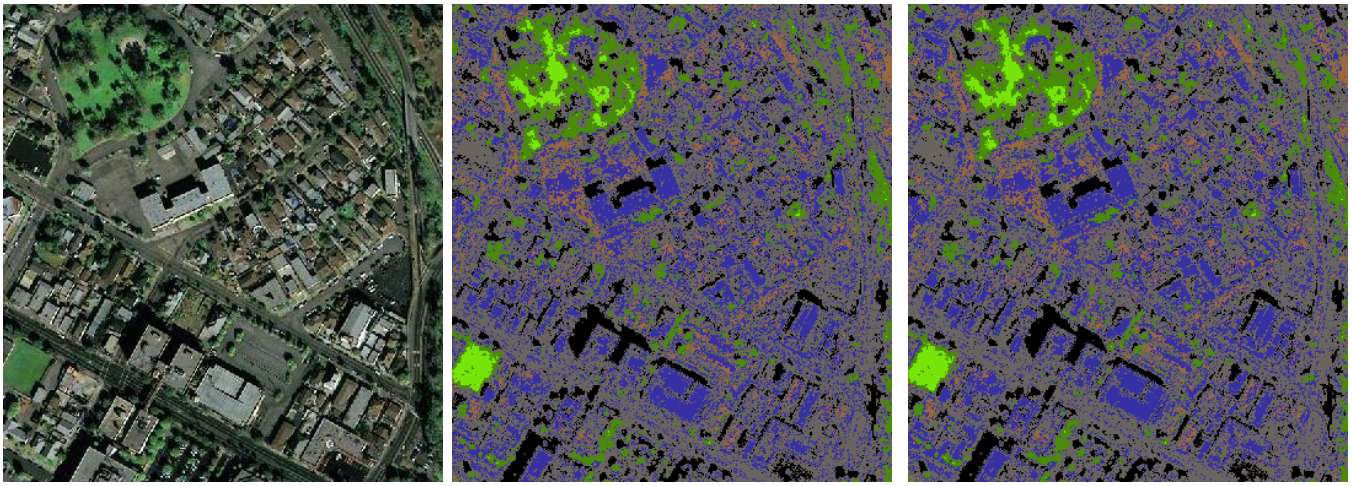
Fig 6. (a) Original image 1, (b) Classification Map obtained using spectral information and haralick features with EMD , (c) Classification Map obtained using spectral information and haralick features with SAM

(a)



(b) (c)

Fig 7. (a) Original image 2, (b) Classification Map obtained using spectral information and haralick features with EMD , (c) Classification Map obtained using spectral information and haralick features with SAM



(a) (b) (c)

Fig 8. (a) Original image 3, (b) Classification Map obtained using spectral information and haralick features with EMD , (c) Classification Map obtained using spectral information and haralick features with SAM

Legend of fig 6, 7 and 8

Asphalt	Green area	Tree	Soil	Building	Shadow

7. ACKNOWLEDGMENTS

This work was funded by CNRST Morocco and CNRS France Grant under "Convention CNRST CRNS" program SPI09/12.

8. REFERENCES

[1] Samson C. 2000. Contribution à la classification des images satellitaires par approche variationnelle et équations aux dérivées partielles : Thesis of doctorate, university of Nice-Sophia Antipolis.

[2] Townshend, J.R.G., 1992. Land cover. International Journal of Remote Sensing 13:1319–1328.

[3] Hall, F.G., Townshend, J.R., Engman, E.T., 1995. Status of remote sensing algorithms for estimation of land surface state parameters. Remote Sensing of Environment 51:138–156.

[4] Lu, D.,Weng, Q., 2007. A survey of image classification methods and techniques for improving classification performance. International Journal of Remote Sensing 28:823–870.

[5] Huang, C., Davis, L.S., and Townshed, J.R.G., 2002. An assessment of support vector machines for land cover classification. International Journal of Remote Sensing 23:725–749.

[6] Kavzoglu, T., Reis, S., 2008. Performance analysis of maximum likelihood and artificial neural network classifiers for training sets with mixed pixels. GIScience and Remote Sensing 45:330–342.

[7] Pal, M., and Mather, P. M. 2005. Support vector machines for classification in remote sensing. International Journal of Remote Sensing, 26:1007−1011.

[8] Zhu, G., and Blumberg, D. G. 2002. Classification using ASTER data and SVM algorithms: The case study of Beer Sheva, Israel. Remote Sensing of Environment, 80:233-240.

[9] Scholkopf, B., Sung, K., Burges, C., Girosi, F., Niyogi, P., Poggio, T., et al. 1997. Comparing support vector machines with gaussian kernels to radial basis function classifiers. IEEE Transactions on Signal Processing, 45:2758−2765.

[10] Cao X., Chen J., Imura H., Higashi O., 2009. A SVM-based method to extract urban areas from DMSP-OLS and SPOT VGT data. Remote Sensing of Environment 113:2205–2209.

[11] Inglada J., 2007. Automatic recognition of man-made objects in high resolution optical remote sensing images by SVM classification of geometric image features. ISPRS Journal of Photogrammetry & Remote Sensing 62: 236–248.

[12] Kruse, F.A., Lekoff, A.B., Boardman, J.W., Heidebrecht K.B., Shapiro, A.T., Barloon, P.J. and Goetz, A.F.H. 1993. The Spectral Image Processing System (SIPS) – interactive visualization and analysis of imaging spectrometer data. Remote Sensing of Environment, 44 : 145-163.

[13] De Carvalho, O.A. and Meneses, P.R. 2000. Spectral Correlation Mapper (SCM): An Improvement on the Spectral Angle Mapper (SAM). NASA JPL AVIRIS Workshop.

[14] Chang, C.-I. 2000. An information theoretic-based approach to spectral variability, similarity and discriminability for hyperspectral image analysis," IEEE Trans. Inf. Theory 46(5):1927–1932.

[15] Robila, S.A. 2005. Using spectral distances for speedup in hyperspectral image processing. International Journal of Remote Sensing 26(24): 5629-5650.

[16] Angelopoulou, E., Lee, S. and Bajcsy, R. 1999. Spectral gradient: a material descriptor invariant to geometry and incident illumination. The Seventh IEEE International Conference on Computer Vision, IEEE Computer Society, 861–867.

[17] Keshava, N., 2004. Distance metrics and band selection in hyperspectral processing with applications to material identification and spectral libraries. IEEE Transactions on Geoscience and Remote Sensing 42: 1552–1565.

[18] Wang, Z., Hu, G., Zhou, Y. and Liu, X. 2008. A classification model of Hyperion image base on SAM combined decision tree. Proc. of SPIE 7146, 71461W.

[19] Bekkari A., Idbraim S., Mammass D. and El yassa M. 2011. Exploiting spectral and space information in classification of high resolution urban satellites images using Haralick features and SVM. IEEE 2ed International Conference on Multimedia Computing and Systems ICMCS'11 , Ouarzazate, Morocco.

[20] Fauvel M., Benediktsson, J. A., Chanussot J. and Sveinsson, J. R., 2007. Spectral and Spatial Classification of Hyperspectral Data Using SVMs and Morphological Profiles. IEEE International Geoscience and Remote Sensing Symposium, IGARSS 07, Barcelona Spain.

[21] Haralick, R. 1979. Statistical and structural approaches to texture. Proceedings of the IEEE 67:786–804

[22] Ohanian, P., Dubes, R. 1992. Performance evaluation for four classes of textural features. Pattern Recognition 25 :819–833

[23] Chiu, W. Y., and Couloigner I. 2004. Evaluation of incorporating texture into wetland mapping from multispectral images. University of Calgary, Department of Geomatics Engineering, Calgary, Canada, EARSeL eProceedings.

[24] Gotlieb, C.C., Kreyszig, H.E. 1990. Texture descriptors based on co-occurrence matrices. Computer Vision, Graphics and Image Processing 51:70–86

[25] Haralick, R.M., Shanmugam K., and Dinstein I., 1973. Textural Features for Image Classification. IEEE Transactions on Systems Man and Cybernetics.

[26] Weszka, J.S., Dyer, C.R. and Rosenfeld, A. 1976. A Comparative Study of Texture measures for Terrain Classification. IEEE Transactions on Systems Man and Cybernetics.

[27] Conners, R.W. and Harlow, C.A., 1980. A Theoretical Comaprison of Texture Algorithms. IEEE Transactions on Pattern Analysis and Machine Intelligence.

[28] Conners, R.W., Trivedi, M.M., and Harlow, C.A., 1984. Segmentation of a High-Resolution Urban Scene using Texture Operators. Computer Vision, Graphics and Image Processing.

[29] Arvis V.; Debain C.; Berducat M.; Benassi A. 2004). Generalization of the cooccurrence matrix for colour images: application to colour texture classification. Journal Image Analysis and Stereology, 23:63-72.

[30] Chapel L. 2007. Maintenir la viabilité ou la résilience d'un système : les machines à vecteurs de support pour rompre la malédiction de la dimensionnalité ?. Thesis of doctorate, university of Blaise Pascal - Clermont II.

[31] Aseervatham S. 2007. Apprentissage à base de Noyaux Sémantiques pour le traitement de données textuelles. Thesis of doctorate, university of Paris 13 –Galilée Institut Laboratory of Data processing of Paris Nord.

[32] Bousquet O., 2001. Introduction au Support Vector Machines (SVM). Center mathematics applied, polytechnique school of Palaiseau. http://www.math.u-psud.fr/~blanchard/gtsvm/index.html.

[33] Keshava, N. and Boettcher, P., 2001. On the relationships between physical phenomena, distance metrics, and best bands algorithms in hyperspectral processing. Proc. Of SPIE, 4381:55–67.

[34] Fauvel M., Chanussot J. and Benediktsson J. A. 2006. A Combined Support Vector Machines Classification Based on Decision Fusion. IEEE International Geoscience and Remote Sensing Symposium, IGARSS 06, Denver, USA.

[35] Hsu, C.W., Chang, C.C., Lin, C.J., 2008. A practical guide to support vector classification.

http://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf .

[36] Chen, S.T., Yu, P.S., 2007. Real-time probabilistic forecasting of flood stages. Journal of Hydrology 340:63–77.

[37] SVMlight Version: 6.02 2008. Developed at University of Dortmund, Informatik, AI-Unit Collaborative Research Center on 'Complexity Reduction in Multivariate Data' (SFB475). http://svmlight.joachims.org/

[38] Scholkopf, B. and Smola, A. J. 2002. Learning with Kernels, MIT Press.