Comparative analysis of Support Vector Machines and Mahalanobis algorithms for road extraction from high resolution satellite imagery

Alberto García Gómez[1], David Fernández Álvarez[2], José Antonio Malpica Velasco[3]

Mathematics Department. Alcalá University. Campus Universitario
28871 Madrid, Spain
[1]j1902@hotmail.com
[2]sendatiempo@hotmail.com
[3]josea.malpica@uah.es

Abstract. High resolution satellite imagery has great potential for a wide variety of cartographic applications, one of the most important being map updating. This has made the automatic extraction of cartographic objects, especially roads and buildings, a challenge in remote sensing. Currently, road extraction from satellite imagery is mostly performed manually; this is a time-consuming task prone to human error. Therefore, automatic or semi-automatic extraction methods are desirable.

In this paper two algorithms, Support Vector Machines (SVM) and Mahalanobis distance have been compared for supervised classification of road networks over high resolution imagery. The segmentation of the satellite image in two classes, road and background, is the first step in most automatic road extraction methods.

Ikonos and SPOT 5 imagery of Alcala, Spain, have been used to test the two algorithms. Confusion matrices and the Receiver Operator Characteristic (ROC) have been used for evaluation of the algorithms. Results show the superiority of the SVM algorithm over the Mahalanobis distance. Another important achievement is that only a small number of training pixels (100 for the imageries studied) are necessary for a classification with an accuracy of 90%.

Furthermore, LIDAR data have been added to the imagery in order to improve the road extraction. Many false positives in the original algorithms are produced due to asphalt roofs; these false positives are removed when LIDAR is considered in the classification process.

Keywords: SVM, Mahalanobis, LIDAR, SPOT 5, road extraction, comparative.

1 Introduction

Today, through the use of satellites, large volumes of data are obtained that are invaluable in the study of Earth's resources and the effects of human activities. However, from a cartographic point of view, remote sensing has focused on land-use and land-cover classification.
Recently, there has been considerable progress in digital photogrammetry. Digital Elevation Models (DEM) and digital orthophotos can now be generated automatically. However, the identification and mapping tasks are still done manually.
Countless public and private companies require current and complete digital map data. It is difficult to keep traditional map products updated because of the rapid pace of development in which we are living. It is therefore necessary to develop a quick and effective method for classifying satellite imagery mapping information.
This paper has two main purposes: first, to find which of the two methods of supervised classification, Support Vector Machines (SVM) or Mahalanobis distance, provides a better identification for roads and second, to establish a methodology to attain high hit ratios with lower training.

2 Datasets and methodology

2.1 SPOT 5 2006

The SPOT 5 image 2006 was taken at 11:23 GMT on August 6, 2006, with an incidence angle of 20.05° and an azimuth of 16.38°. The sun had an elevation of 62.37º and an azimuth of 150.06°.

Launch Date	May 3, 2002
Launch Vehicle	Ariane 4
Launch Location	Guiana Space Centre, Kourou, French Guyana
Orbital Altitude	822 kilometers
Orbital Inclination	98.7°, sun synchronous
Speed	7.4 km/s (26,640 km/h)
Equator Crossing Time	10:30 AM (descending node)
Orbit Time	101.4 min
Revisit Time	2–3 days, depending on latitude
Swath Width	60 km x 60 km to 80 km at nadir
Metric Accuracy	< 50 m horizontal position accuracy (CE 90%)
Onboard Data Recorder (Gb)	90
Data Rate (Mbps)	100
Number of Sensors	4

Table 1 General Characteristics of the SPOT 5 Satellite

2.2 SPOT 5 2006+ LIDAR

Four SPOT 5 bands were combined with LIDAR data in order to check a possible improvement in the classification of roads. LIDAR data were derived as a subtraction of the digital surface model (DSM) from the digital terrain model (DTM) and in that way the normalized digital surface model (nDSM) (Fig. 1) was obtained.

Fig. 1 Normalized Digital Surface Model (nDSM = DSM-DTM)

2.3 Support Vector Machines

The algorithm SVM is formed by a group of supervised learning methods that can be applied to the classification and regression. From a theoretical point of view, it is based on a statistical learning theory proposed by Vapnik and Chervonenkis [1].

2.4 Mahalanobis distance

Assuming that there is a set of training pixels and each class has a Gaussian distribution after several mathematical operations [2], the formula for the Mahalanobis distance is:

$$\delta(x) = (x - \vec{m})^T \sum^{-1} (x - \vec{m}) \quad (1)$$

After calculating the Mahalanobis distance for each pixel of the image, we normalize the Mahalanobis distance to a range of values [0,1] to be compared with the results produced by the SVM. Thus, a pixel with a Mahalanobis distance of 0 will be assigned a probability of 1 of belonging to that class. The higher the distance, the lower the probability.

3 DESCRIPTION AND METHODOLOGICAL DIAGRAM

3.1 Phase I. Correction and Classification of images

A preliminary step is to merge multispectral and panchromatic bands, using a method called pansharpening, a fusion technique used to improve the spatial resolution of the multispectral images.
With the images resulting from the merger, an unsupervised classification is performed to observe the distribution of the pixels in different classes according to their spectral characteristics. This provides a better understanding of the distribution of the classes of the image that will be useful when performing the supervised classification.
Based on the analysis of the statistics obtained, we get the training areas for the most representative background class. These training areas are selected following a strict proportionality test according to their degree of presence in the picture. Later, we proceed to the election of the samples (pixels) of roads for training for the supervised classifications.
We use different training levels: high, medium, and low (10.000, 6.000, and 1.000 pixels, respectively) for the study, while the background remains constant (17.000 pixels) mainly for two reasons: first because the variation is likely due to different types of roads and second, because road cover has areas with very heterogeneous spectral characteristics. Thus we establish a methodology to train the image. It is important to clarify that the classification is performed with a training of 10% of the total pixels in the training sample; we then validate with the 90% remaining pixels.
In the next section we discuss the supervised classification.

3.2 Phase II. Study of the classifications and possible improvement with LIDAR

Once we obtain the classified images, we start the second phase where the evaluation methods shown below are applied. From this we establish which study provides the best method. Only the visual study of thematic maps cannot establish which of the two classifiers is more effective. We use a variety of statistical analysis tools: confusion matrix, kappa coefficient, Jeffries–Matusita distance, and ROC area. Finally, the data obtained using these tools will be the ones to determine the applicability of each algorithm.
Two images are studied: SPOT 5 from 2006 and the same SPOT 5 built with an extra band (nDSM) extracted from LIDAR.

4 DISCUSSION AND RESULTS

As explained above, we work with four tools to evaluate which, SVM or Mahalanobis, is the most effective method to classify roads.
Now we see how the values of these tools vary depending on the level of training and the type of image. A series of graphs that analyze the two classification methods, showing the variations with different degrees of training on the road (low, medium, and high) and a constant background in 17.000 pixels was used to show the classification results.

4.1 Jeffries–Matusita distance

The Jeffries–Matusita (JM) distance is an indicator of how classes are separated. It comes from the Bhattacharyya distance, given by

$$B=\frac{1}{8}\ \mu_2-\mu_1\ ^{t}\left(\frac{\Sigma_1+\Sigma_2}{2}\right)^{-1}\ \mu_2-\mu_1\ +\frac{1}{2}\log\left(\frac{\left|\Sigma_1+\Sigma_2\right|/2}{\sqrt{\left|\Sigma_1\right|\left|\Sigma_2\right|}}\right), \quad (2)$$

where $\mu_i\,\Sigma_I$, i=1,2 are the median and covariate matrices for classes 1 an 2.
Large values of B imply good reparability. The first term in B is a squared average Mahalanobis distance. The JM distance measures separability of two classes on a more convenient scale in terms of B.

$$JM=2\ \ 1-e^{-B} \quad (3)$$

The value 2 is the highest and 0 the minimum separability. In principle this means that when regions of interest are better defined, the likelihood of being classified correctly is higher.

4.2 Receiver Operator Characteristic

This method gives the area enclosed by a curve facing the likelihood of false positives with the probability of detection of a class. We must also take into consideration that the form that describes the curve is equally important.
For the 2006 SPOT 5 image, the SVM classifier is always above Mahalanobis classifier (Fig. 3), while when considering SPOT 5 plus LIDAR, the SVM classifier is better only with a high training (Fig. 4), meaning many pixels in the training samples. It can be observed that the curve describing Mahalanobis is irregular and has abrupt changes, while SVM is more regular and smooth (Fig. 2). Overall SVM is superior to Mahalanobis.

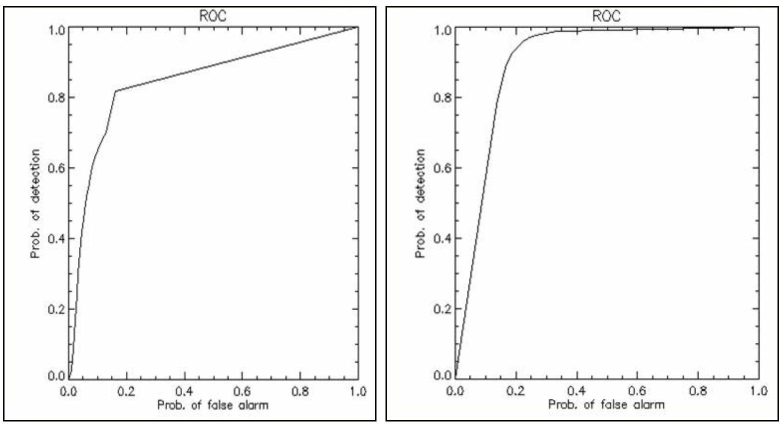


Fig. 2 ROC area: Mahalanobis (left) and SVM (right)

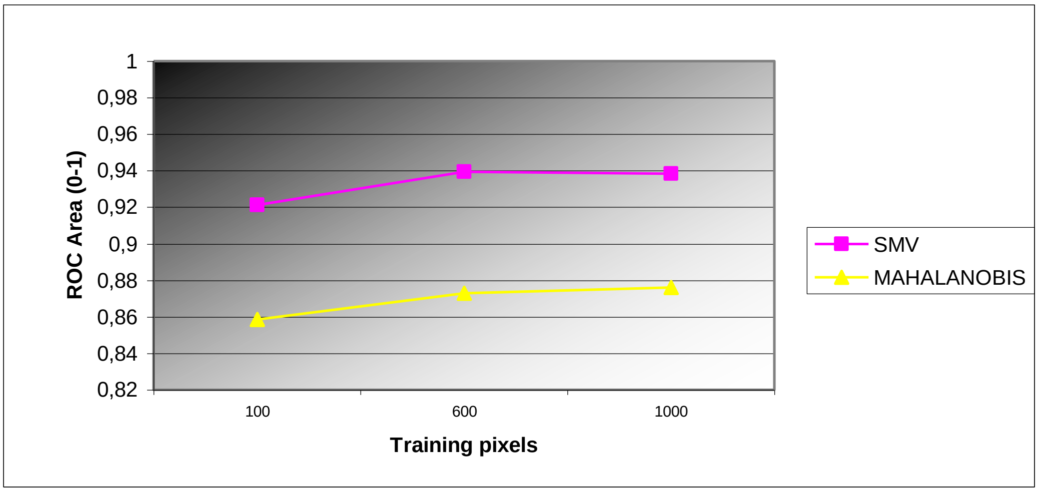


Fig. 3 ROC area in SPOT 5 (2006)

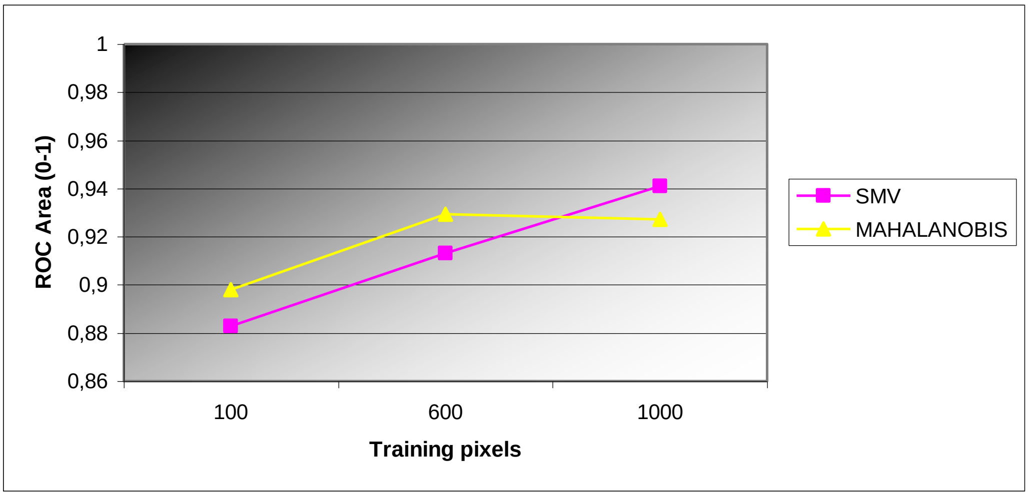


Fig. 4 ROC area with SPOT 5 + LIDAR

4.3 Confusion matrix

In the field of artificial intelligence, confusion matrix is a visualization tool that is used in supervised learning and provides the percentage of well-classified elements on the total sample.
However, if the classes are unbalanced, this method can produce false results. If, for example, there are 990 samples of Class 1 and only 10 of Class 2, the classifier can easily tend toward Class 1. If the classifier identified all samples as Class 1, accuracy is 99%. This does not mean it is a good method of classification as it has 100% error in the classification of samples of Class 2. That is why the confusion matrix should be complemented with the "kappa coefficient," which is explained below.
Examining the graphs we conclude that SVM is always higher than Mahalanobis, reaching 100% in any image and at any level of training. In summary, both with low, medium, or high training, SVM obtains a better success rate than Mahalanobis (Figs. 5 and 6).

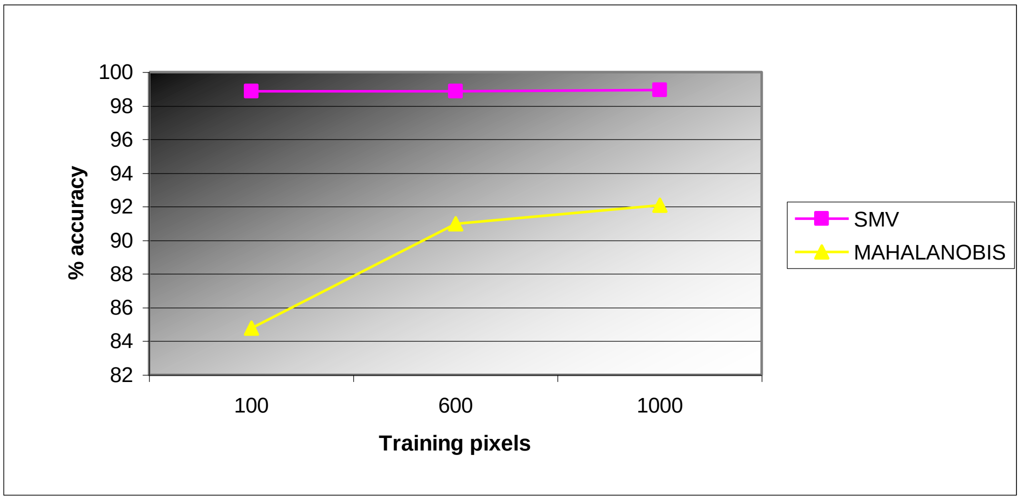


Fig. 5 Confusion matrix in SPOT 5 (2006)

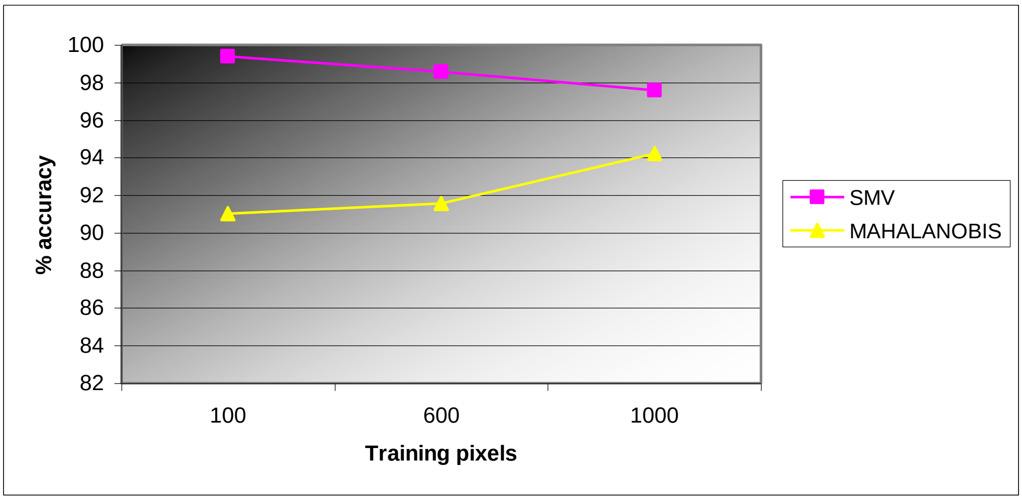


Fig. 6 Confusion matrix with SPOT 5 + LIDAR

4.4 Kappa Coefficient

As we saw in the previous paragraph, the confusion matrix needs to be complemented by the kappa coefficient. In summary, we conclude that when the kappa coefficient is closest to 1, the classification will be better. In the SPOT 5 image, SVM values are always higher than those of Mahalanobis, exceeding 85% success with low training (Fig. 7). In the second image with the extra band of LIDAR and low training, Mahalanobis is a bit better, but with more training the trend changes (Fig. 8).

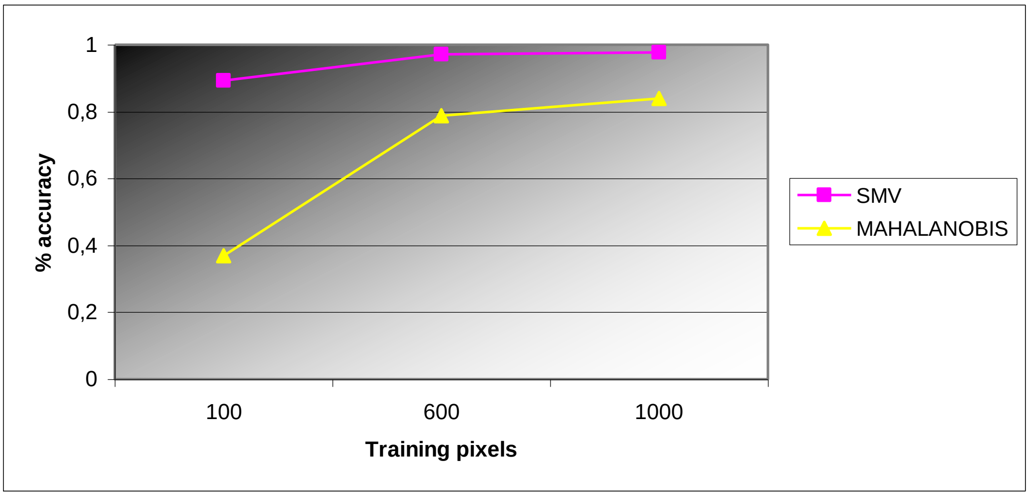


Fig. 7 Kappa coefficient in SPOT 5 (2006)

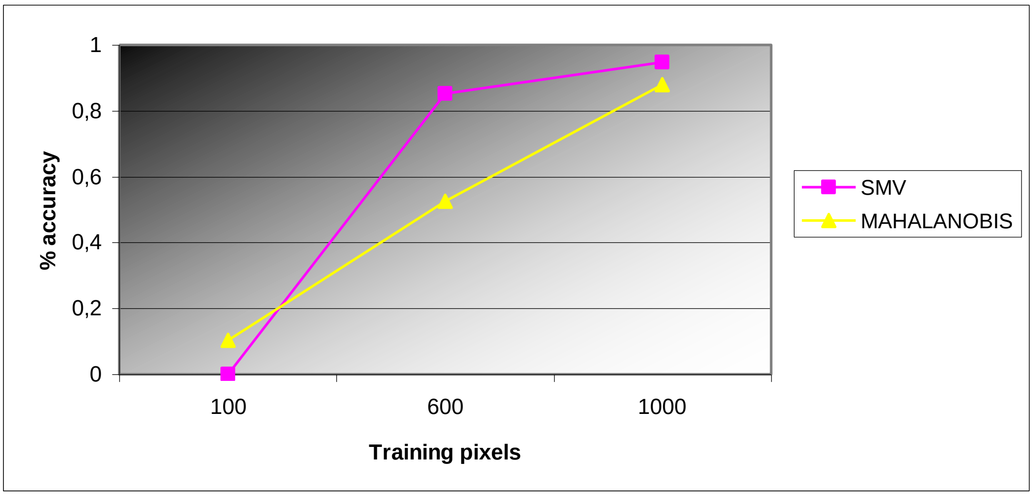
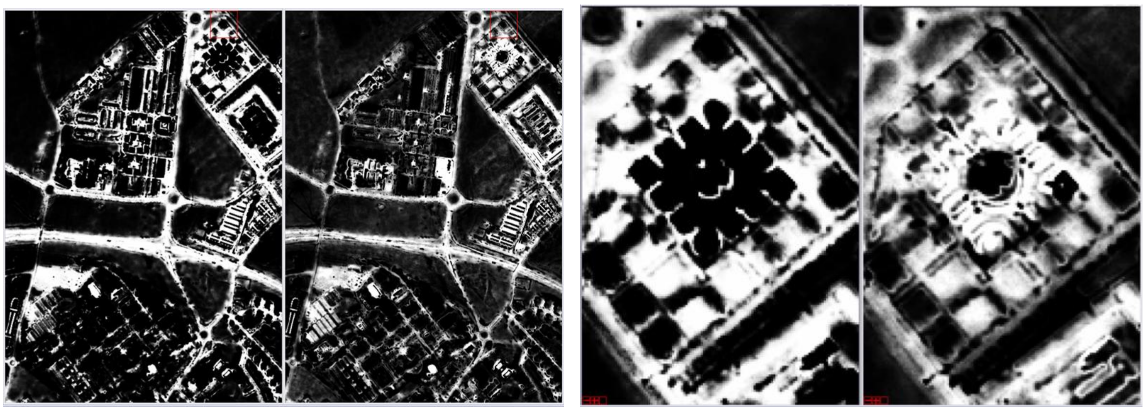


Fig. 8 Kappa coefficient in SPOT 5 + LIDAR

5 CONCLUSIONS

Studying the obtained graphs, it is not difficult to conclude that SVM is clearly superior to Mahalanobis, at least as far as detection of paved roads is concerned. If we follow the evolution through the various levels of training, we see that SVM gives a greater area in the ROC curves, a better ratio of hits in the confusion matrix, and is more precise classifying the pixels in their true classes than Mahalanobis.
Analyzing the results, the level of training of roads that should be taken to obtain a rating of at least 90% of well-classified pixels would be 1.000 (note that the classification is made with 10% (100 pixels) and is validated with the remaining 90%). Descending from these levels, the classification is more irregular and inefficient.
On the other hand we see that adding LIDAR does not dramatically improve the statistics, but we obtain a more accurate classification, eliminating false positives that appear as a consequence of asphalt roofs of buildings, as seen in Figure 9.



a b c d

Fig. 9 Effects of adding LIDAR data to the multispectral satellite data. This images are probability layers for the classification of the four multispectral Spot 5 2006 satellite bands without and with LIDAR. Classification with LiDAR (a), classification with out LIDAR (b), detected roof of building when adding LIDAR (c) undetected roof when LIDAR is not considered (d).

Acknowledgments

The authors wish to thank the Spanish National Mapping Agency (Instituto Geográfico Nacional) for providing the images for this study, and to the Spanish Ministry of Science and Innovation (project CGL2009-13768) for financial support for presenting this paper. Thanks are also due to the Ostrava-GIS reviewer who helped to improve this paper.

Reference

[1] Christopher J.C. Burges, 1998. A Tutorial on Support Vector Machines for Pattern Recognition. Data Mining and Knowledge Discovery (2) pp. 121-167.

[2] Richards, John A., Xiuping Jia, 2006. Remote Sensing Digital Image Analysis: An Introduction, 4th edn. 439 pages. Springer, Heildelberg.

[3] Swets, J.A, 1979. ROC analysis applied to the evaluation of medical imaging techniques, Investigate Radiology, 14: 109-121.

[4] Vapnik, V., 1995 The Nature of Statistical Learning Theory. Springer, New York.