IDENTIFICATION OF VEGETATION CHANGES USING
BI-TEMPORAL SPOT 5 IMAGES

Jose A. Malpica
Maria C. Alonso
Mathematics Department
University of Alcala
28871 Madrid, Spain
josea.malpica@uah.es
mconcepcion.alonso@uah.es

ABSTRACT

Our main objective is to present a methodology for vegetation change detection in multi-temporal, multi-spectral satellite images. The method consists of applying a classifier to each image in order to obtain the probability layers for vegetation, and then analyzing the differences to detect the changes. The method was evaluated using two high-resolution SPOT 5 satellite images, taken during the successive summers of 2005 and 2006, with a similar azimuth and elevation of the sun.

Two supervised algorithms have been tested: Support Vector Machine (SVM) and Mahalanobis distance. The SVM classifier is compared with the Mahalanobis classifier. The outcome is that the SVM results in an image that more clearly shows changes. The evaluation is accomplished visually, using ROC curves. The NDVI index is also considered, since it is the usual means of studying vegetation in satellite images. The NDVI is clearly inferior to the SVM and Mahalanobis classifiers in terms of accuracy, although the latter two need require an improved training set, which could be a disadvantage for some applications.

INTRODUCTION

Change detection consists of identifying environmental changes by observing a geographic area over a period of time. This paper focus on detecting changes in green vegetation by means of remote sensing. Change detection can be performed by comparing an image taken at a specific date with an older map, or comparing two images taken on different dates. Here, we will deal only with the latter case, analyzing bi-temporal images from the same SPOT 5 sensor. It is recommended to work not only with the same sensor, but also with imagery taken at the same epoch of the year, in order to avoid or reduce the effect of seasonal differences and sun angle incidence. However, even when special care is taken to collect the best data possible, it is difficult to avoid some radiometric differences between the images. Many papers about change detection have been published in the last three decades. Singh (1989) provided an early review of the literature, followed more recently by Lu et al. (2003) and Coppin et al. (2004).

Low-resolution satellite imagery commonly uses pixel-by-pixel change detection methods that cannot generally be applied to high-resolution satellite imagery, such as SPOT 5. This is because SPOT 5 images have greater detail and are thus more complex. These high-resolution images contain many new signatures from different materials that do not appear in low-resolution images. We will try to work around this problem, using the probability layer instead of working pixel-by-pixel or with completed classified images.

The ability to detect detailed vegetation change by using high-resolution remotely sensed image data is of utmost importance to many environmental and agricultural projects (Jensen, 2007; Lillesand, 2008). The quantitative assessment of biomass change over time is of great importance in monitoring deforestation or vegetation phenology. In situ data acquisition is costly in terms of both effort and money, and when the study area is large the data can be inaccurate. Remote sensing provides a useful method for assessing biomass flux over time.
MATERIALS AND METHODS

Datasets

The images we used in our experiment come from SPOT 5 high spatial resolution sensors. The SPOT 5 satellite was launched on May 2, 2002 and is the fifth satellite in the SPOT series. SPOT 5 uses sophisticated processing techniques to obtain a 2.5 meter resolution from two 5-meter images acquired simultaneously by the sensor with two rows of 12,000 CCD detectors. This allows optimum image quality in terms of resolution and solar illumination, even though the satellite is about 800 km above Earth. The resolution (see Table 1) for bands B1, B2 and B3 is 10 m, while the resolution for B4 is 20 m. SPOT 5 provides a good balance between high resolution and wide-area coverage (60 km x 60 km). The acquired images were of an area near Madrid, Spain that measured 20 x 13 km. The first image is a SPOT 5 subscene taken in the morning (11:09 hours) of 24 July, 2005 with an incidence angle of 3.12 and an orientation angle of 14.01. The sun had an elevation of 44.23 degrees and an azimuth of 166.69 degrees (Figure 1 (a)). The second image is also a SPOT 5 subscene, taken in the morning (11:23 hours) of 6 August, 2006 with a satellite incidence angle of 20.05 left and an orientation angle of 16.38. The sun had an elevation of 62.37 degrees and an azimuth of 150.06 degrees (Figure 1(b)).

![Images](a) ![Images](b)

Figure 1. SPOT-5 represented in false color with bands 1, 2 and 3 for RGB, 2005 scene (a) and 2006 scene (b).

Preprocessing

The images were first georeferenced to the European Datum with precisions better than one pixel; therefore no co-registration was necessary. Pan-sharpening using PCA transformation was applied to both images; thus, two sets of four multi-spectral images were obtained with a resolution of 2.5 m.

Radiometric correction due to atmospheric conditions is difficult because it is necessary to know the exact conditions at the time the images were taken, and this information is not always available. Radiometric normalization is also difficult because there are always different reflections and tonality in corresponding pixels due to the different altitudes of the sun and the different incident angles in the two images. However, atmospheric correction or normalization is not necessary in this case because the proposed method will calculate a probability layer independently for each image. As Singh (1989) stated, when images are classified independently, the effects due to different atmospheric conditions and different sensors for multispectral image acquisition are minimized.

Classification Algorithms

We are going to apply the Normalized Difference Vegetation Index (NDVI) (Jensen, 2007) for vegetation extraction as well as two supervised algorithms, the Support Vector Machine (SVM) and the Mahalanobis distance. For the sake of clarity, a brief description of the NDVI index and the two algorithms used for the classification are provided. Refer to Vapnik (1995) for more detail about the Support Vector Machine (SVM), and to Richards (2006) for more detail about Mahalanobis distance.

---

Table 1. SPOT 5 characteristic

<table>
<thead>
<tr>
<th>Features</th>
<th>SPOT 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spectral bands</td>
<td>P : 0,49–0,69 µm (Panchromatic, 2.5m)</td>
</tr>
<tr>
<td></td>
<td>B1 : 0,50–0,59 µm (Green, 10m)</td>
</tr>
<tr>
<td></td>
<td>B2 : 0,61–0,68 µm (Red, 10m)</td>
</tr>
<tr>
<td></td>
<td>B3 : 0,79–0,89 µm (Near-infrared, 10m)</td>
</tr>
<tr>
<td></td>
<td>B4:1,58–1,75µm(Middle-Infrared, 20m)</td>
</tr>
<tr>
<td>Sensor footprint</td>
<td>60 km x 60 km</td>
</tr>
<tr>
<td>Revisit interval</td>
<td>Every 26 days</td>
</tr>
</tbody>
</table>

---

ASPRS 2009 Annual Conference
Baltimore, Maryland • March 9-13, 2009
1. NVDI:
The NVDI is a simple arithmetical operation of the image bands

\[
NDVI = \frac{\text{NIR} - \text{VIS}}{\text{NIR} + \text{VIS}}
\]

(1)

NIR and VIS stand for near infrared and visible bands, respectively. For a given pixel, the calculation of the NDVI ranges from minus one to plus one and can be normalized to the range 0 to 255 to be represented as an image.

2. SVM:
The idea for SVM initially appeared in an article by Boser et al. (1992) in which they applied SVM to solve optical character recognition problems. They demonstrated the superior generalization of SVM compared with other learning algorithms. SVM maximizes the margin between the training patterns and the decision boundary.

SVM is basically a binary classifier by nature; however, it can be redesigned to handle the multiple classification problems that commonly arise in remote sensing applications. The two approaches commonly used to do this are the One-Against-All and the One-Against-One techniques. To date, there has been no research stating that one method is superior to the other (Hao, 2006, X).

A pixel is represented with the vector \( x \), with 4 components corresponding to the images of the four pansharpening bands.

Let us take two classes represented by: \( \omega_i, i = 1,2 \). The first class corresponds to the object of interest for the analysis of its changes. For example, this could be green vegetation, i.e., plant life with chlorophyll that receives a red color on the infrared SPOT 5 band (see Figure 1). The other class, \( \omega_2 \), is considered to be the background.

To illustrate the algorithm, let us start with a two class problem. Given some training data by a set of points of the form

\[
T = \{ (x_i, z_i) | x_i \in \mathbb{R}^n, z_i \in \{-1,1\} \} \quad i = 1 \ldots m,
\]

(2)

where \( x_i \) is a point or an n-dimensional feature vector of the training set, which has cardinality \( m \), and \( z_i \) is a label indicating the class to which the element \( x_i \) belongs, either 1 or -1 (as mentioned above, this is a two class problem).

The set of points \( x \) satisfying the equation

\[
g(x) = w \cdot x + b \leq 0,
\]

(3)

form a hyperplane of \( \mathbb{R}^n \), where the vector \( w \) is perpendicular to the hyperplane, and \( \frac{b}{||w||} \) represents the offset of the hyperplane form the origin along the vector \( w \). All of the training points will lie on one or the other side of the hyperplane \( g(x) \);

\[
\begin{cases}
(w \cdot x_i + b) > 0 & \text{if } z_i = 1 \\
(w \cdot x_i + b) < 0 & \text{if } z_i = -1.
\end{cases}
\]

(4)

There are infinite hyperplanes that separate the two classes of the training points. Let us select the hyperplane that maximizes the perpendicular distance to the points nearest to the hyperplane on both sides.

Scaling \( w \) by a constant and fixing \( b \) appropriately, the following equations for the parallel bounding hyperplanes can be obtained

\[
\begin{cases}
(w \cdot x_i + b) \geq 1 & \text{if } z_i = 1 \\
(w \cdot x_i + b) \leq -1 & \text{if } z_i = -1.
\end{cases}
\]

(5)

In order to express this as only one equation, equation 4 can be rewritten as:
\[ z_i (w \cdot x_i + b) \geq 1. \]  

\[ \begin{align*}  
&g(x) = 1 \\
g(x) = 0 \\
g(x) = -1 \\
support vectors
\end{align*} \]

Figure 2. Support Vector Machine for linearly separable data.

The objective is to find the separating hyperplanes with the largest gap or margin. This objective is accomplished when we find the weight vector \( w \) that maximizes the distance between the \( g(x) = 1 \) and \( g(x) = -1 \) hyperplanes. Basic geometric calculation tells us that the distance between these hyperplanes is \( \frac{2}{\|w\|} \). Therefore, maximizing the hyperplane margin will be the same as minimizing \( \|w\| \), i.e.

\[
\text{minimize: } \frac{1}{2} \|w\|^2, \quad \text{subject to } z_i (w \cdot x_i + b) \geq 1 \quad \forall i, \quad i = 1..n. \tag{7}
\]

This can be now resolved through quadratic programming techniques. The support vectors are the training samples that define the optimal separating hyperplanes, i.e. the training elements for which equation 4 becomes equality (Figure 2). By the Lagrange method of undetermined multipliers, the problem expressed in equation (7) can be reformulated as:

\[
f(w, \lambda_k) = \frac{1}{2} \|w\|^2 - \sum_{k=1}^{n} \lambda_k (z_k w^T x_k - 1). \tag{8}
\]

The generalization to a multiclass problem is achieved by calculating \( c \), linear discriminant functions. Mercier and Lennon (2003) reported that high classification accuracy can be accomplished with small training sets. SVM provides a mathematical foundation for how well the classifier will generalize to unseen data. Furthermore, the SVM training algorithm is guaranteed to converge to the globally optimal SVM classifier and can learn highly non-linear discrimination functions.

Several researchers have applied SVM to images. Azimi-Sadjadi and Zekavat (2000) used SVM to classify ten different cloud-covered and cloudless areas. SVM has been compared to other classification methods for remote sensing imagery, such as Neural Networks, Nearest Neighbor, Maximum Likelihood and Decision Tree classifiers, and has surpassed them all in robustness and accuracy (Huang et al. 2002, Foody and Mathur 2004, Melgani and Bruzzone 2004, Theodoridis and Koutroumbas 2003). These results encouraged us to apply SVM and we did not find big differences in the Mahalanobis distance (as we will see below) related to the Maximum Likelihood.

3. Mahalanobis
In order to know to which class \( \omega_i \) belongs, a given pixel, \( x \) should be calculated for the following probabilities.

\[ P(\omega_i \mid x), \ i = 1,2. \tag{9} \]

A pixel is assigned to the object of interest if its conditional probability is the greatest.
\begin{equation}
  x \in \omega_i \iff P(\omega_i \mid x) > P(\omega_j \mid x) \quad \forall j \neq i. \tag{10}
\end{equation}

Next, we assume there is a training set of pixels. Assuming the distribution to be Gaussian for each class and following Richards and Jia (2005), after several mathematical operations everything converges to the popular Mahalanobis distance as the classifier,

\begin{equation}
  \delta(x) = (x - m)^T \sum^{-1} (x - m) \tag{11}
\end{equation}

The Mahalanobis distance for each pixel is calculated for the SPOT 5 scene, then the whole range of values is normalized to the interval [0,1]. If a pixel has a Mahalanobis distance of 0, it will be assigned a probability of 1; this means that the pixel belongs to the class \( \omega_1 \). The opposite is also true; if a pixel has the largest Mahalanobis distance, then it will have 0 probability of belonging to class \( \omega_1 \). The larger the distance, the smaller the probability will be.

**Training**

Due to the displacement toward the infrared region of the SPOT 5 sensor, green vegetation is seen with a red color. This is convenient to differentiate green vegetation from other objects, then training samples from areas with green vegetation in both images, where this vegetation has not changed in the two images. In Figure 3 can be seen the training sample draw on the SPOT 2005 and detail.

![Figure 3](image-url)

**Figure 3.** Training samples and details.
RESULTS AND DISCUSSION

The most widely accepted means of estimating biomass with remote sensing is the NDVI index. Figures 4(a) and 4(b) show the NVDI layer for the years 2005 and 2006, respectively.

![Figure 4](image)

*Figure 4.* Different result images: NVDI for 2005 (a), NVDI for 2006 (b), SVM for 2005 (c), SVM for 2006 (d), Mahalanobis for 2005 (e), Mahalanobis for 2006 (f), Difference for SVM (g) and Difference for Mahalanobis.

The NVDI has taken as vegetated areas (brighter pixels) some that are not, such as the synthetic playgrounds in the upper part of the area under study. These types of errors do not occur using the SVM algorithm or the Mahalanobis distance. Looking at the results in Figures 4 (a), (b), (e) and (f), it is obvious that the SVM is superior to the NVDI. Similarly, it is obvious by looking at Figures 4 (a), (b), (e) and (f) that the Mahalanobis distance is also superior to the NVDI. An advantage of the NVDI index, and other similar vegetation indexes, over both the Mahalanobis distance and the SVM is that they do not need any training data. Therefore, it is not fair to compare the NVDI with supervised methods such as the SVM or Mahalanobis distance (even though the SVM does not need many elements in the training data).
It remains to compare the SVM and the Mahalanobis distance. Figures 4(e) and 4(b) show the results of applying Mahalanobis distance to the 2005 and 2006 images. Actually, the SVM and Mahalanobis layers are not direct probability layers, as would be the Maximum Likelihood. However, for all intents and purposes, they could be considered as probability layers; the longer the distance of the pixel to the flat hyperplane in the SVM, the higher the probability that it belongs to the object of interest (in our case, vegetation class) and therefore the brighter the pixel. The same is true when applying Mahalanobis distance. For the Mahalanobis distance, the smaller the Mahalanobis distance, the higher the probability, as explained in the previous section.

Figure 4(g) shows the difference between the SVM layers in 2006 and 2005. Therefore, the brighter pixels in 4(g) show “positive” changes in vegetation, i.e., there was vegetation in 2006 that was not there in 2005. The darker pixels represent “negative” changes, i.e., where there was vegetation in 2005, there was none in 2006. Similarly to the SVM, the difference of layers 4(f) minus 4(e) gives 4(h) for Mahalanobis.

It still remains to discover the differences between 4(g) and 4(h). To discuss this, let us examine detailed images of the results for each classification. Figure 5 shows a detail of the upper right part of the study image corresponding to the campus of the University of Alcalá. One can see a football field that had grass in 2006, but did not in 2005. This is more clearly observable in the SVM image in Figure 5 (c) than in the Mahalanobis image in Figure 5 (d). Furthermore, in the middle left of the image in Figure 5, one can see four circular flowerbeds or parterres. The two in the middle had grass or green vegetation with chlorophyll in 2006, but not in 2005. In fact, these parterres were being built in 2005 (which is why there was no green vegetation). By 2006 they were completed and contained vegetation, as can be seen in Figures 6(a) and 6(b). Figures 6(c) and 6(d) show the results of the classification using the SVM and Mahalanobis distance, respectively. It can be seen how the SVM shows a more clearly defined area of vegetation change. This is because the SVM uses a hyperplane to separate the pixels belonging to green vegetation from those that do not belong to it, while Mahalanobis distance uses the covariance matrix in its definition spreading the discriminate information of the classifier. This will be confirmed by the Receiver Operating Characteristic curves.

Figure 5. Detail of the university campus, 2005 (a), 2006 (b), Difference for SVM (c) and Difference for Mahalanobis (d).
Figure 6. Detail of the parterres, 2005 (a), 2006 (b), Difference for SVM (c) and Difference for Mahalanobis (d).

Visual evaluation is not enough to determine which classifier is better or worse; we need some kind of numerical evaluation. A method that was first applied in medical imaging to evaluate classifiers is Receiver Operating Characteristic (ROC) curves (Swets, 1979), which plot the probability of detection versus the probability of false positives.

The training set in Figure 3 was divided in two sets, using randomization for 10% and for 90% of the pixels, respectively. The two classifiers, SVM and Mahalanobis, were trained with the first set (10% pixels); and both classifiers were tested with the second set (90% of the pixels). We used twenty thresholds (points) for building the ROC curves for SVM and Mahalanobis classifiers, as can be seen in Figure 7.
The accuracy of a classifier can be measured by the area under the ROC curve. An area of 1 represents a perfect test, and an area of 0.5 represents a worthless test. A rough guide for knowing the accuracy of a classifier is: 0.5-0.6: fail; 0.6-0.7: poor; 0.7-0.8: fair; 0.8-0.9: good; and 0.9-1: excellent. The area under the curve for our experiment gives 0.9761 for the SVM and 0.9773 for the Mahalanobis distance. Although the value is slightly better for the Mahalanobis distance, it cannot be said that Mahalanobis distance is better than the SVM. What is interesting is how the curve for the Mahalanobis distance crosses the curve for the SVM. Thus, as can be seen in Figure 7, for a small probability of false alarm Mahalanobis outperforms SVM in probability of detection, while when the probability of false alarm is allowed to have a bigger value, SVM outperform Mahalanobis. This coincides with the intuitive idea about the theoretical fundamentals of each method, and the differences in the way each classifier detected the vegetation changes, The SVM shows a more clearly defined area of vegetation change which corresponds to a steep ROC curve, while Mahalanobis detects a spread defined area of vegetation change which corresponds to a gradual rise of the ROC curve.

CONCLUSIONS

It is common knowledge that different change detection algorithms have their own advantages, and that no single approach can be considered the best for all cases. In practice, different algorithms are often compared to find the optimal change detection algorithm for a specific application. The algorithm proposed in this paper is useful when detecting vegetation changes in high-resolution satellite images similar to SPOT 5. The results prove that the proposed method based on the SVM classifier is more accurate in identifying changes in vegetation than is the usual method of differencing NDVI data. When comparing the SVM and Mahalanobis distance, they are very close in terms of accuracy. However, the former detects vegetation changes more clearly than the latter.

ACKNOWLEDGMENTS

The authors wish to thank the Spanish National Mapping Agency (Instituto Geográfico Nacional) for providing the images for this research. The authors are also grateful to the Spanish Ministry of Education and Science for financial support to present this paper in ASPRS 2009; project number CGL2006-07132/BTE
REFERENCES


