BUILDING DETECTION FROM PAN-SHARPENED IKONOS IMAGERY THROUGH SUPPORT VECTOR MACHINES CLASSIFICATION

M. Turker [a, *] and D. Koc San [b]

[a] Dept. of Geodesy and Photogrammetry, Hacettepe University, 06800 Ankara, Turkey- mturker@hacettepe.edu.tr
[b] Dept. of City and Regional Planning, Selcuk University, 42075 Konya, Turkey; e-mail: dilekoc@gmail.com

Commission VIII, WG VIII/8

KEY WORDS: Building detection, Classification, Support Vector Machines, IKONOS, nDSM, NDVI

ABSTRACT:

An approach is presented for detecting the buildings from high resolution pan-sharpened IKONOS imagery through binary Support Vector Machines (SVM) classification. In addition to original spectral bands, the bands nDSM (normalized Digital Surface Model), NDVI (Normalized Difference Vegetation Index), PC1, PC2, PC3, and PC4 (First, Second, Third, and Fourth Principal Components), are also included in the classification. The proposed classification procedure was carried out in three study areas selected in the Batikent district of Ankara, Turkey. The study areas show different residential and industrial characteristics. The first study area covers mainly the residential parts that include buildings with different shapes, sizes, dwelling types, and colored roofs. The second study area also represents the residential characteristics but contains buildings with more regular shapes. The third study area contains the industrial buildings with the gray tone roofs and the sizes of the buildings are larger. Also tested in the present study is the effect of the training sample size in the accuracy of the SVM classification. The results reveal that the overall accuracies were computed to be between 90% and 99%, while the kappa coefficients were found to be between 0.80 and 0.98. The inclusion of additional bands in the SVM classification had a considerable effect in the accuracy of building detection. Increasing the training size increased the accuracy, however, the increase was not more than 3%.

1. INTRODUCTION

Automatic building extraction from high resolution imagery is a very challenging task and has been the focus of intensive research for the last decade. High resolution satellite images provide a valuable new data source for geographic data acquisition, mapping applications, and urban planning. In the last decade, spaceborne remote sensing has experienced with an intensive technological development. The high resolution satellite images are being increasingly used for the detection of the buildings. Of the techniques used, automatic image classification is the most widely used technique for the detection of the buildings.

In a conventional supervised classification all classes must be defined and trained in detail. If the interest is to a single class, the use of a conventional supervised image classification technique may be inappropriate (Foody et al., 2006). Support Vector Machines (SVM) classification is a supervised classification technique derived from the statistical learning theory. Originally the SVMs were designed for binary classification and then extended for multi-class classification. Foody et al. (2006) claimed that one class SVM classification has a great potential in remote sensing. Boyd et al. (2006) used this classification technique for mapping a specific class for the priority habitats monitoring. Sanchez et al. (2007) used it for mapping a specific class from Landsat 7 ETM+ image.

The past studies have shown that the SVM classification technique provides higher accuracy than the standard parametric and popular classification techniques in the classification of the remote sensing images (Huang et al, 2002; Zhu and Blumberg, 2002; Pal and Mather, 2005; Foody and Mathur, 2004, Watanachaturaporn et al. 2008). Watanachaturaporn et al. (2008) compared SVM classifier with the Maximum Likelihood Classifier (MLC), Backpropagation Neural Network Classifier (BPNN), Radial Basis Function Neural Network Classifier (RBFNET), and the Decision Tree Classifier (DTC) and found that the SVM classifier produced significantly higher accuracy than the others.

Several studies have used multi-class SVM classification for land use detection of urban areas from high resolution satellite images. Tuia et al. (2010) performed SVM classification using composite kernels for the classification of high resolution urban images and concluded that a significant increase in the classification accuracy was achieved when the spatial information was used. Huang et al. (2007) proposed an adaptive multiscale information fusion algorithm to extract the spatial features and classify the high resolution imagery. Li et al. (2010) presented an object-oriented classification method based on the improved colour structure code and SVM. Their results indicate that perfect fusion strategy and classification pre-processing are helpful for improving the classification accuracy. Bellman and Shortis (2000) used SVM Classification technique and Wavelet Transform to detect building areas from aerial images and they concluded that this technique appears to be quite promising for use in the initial building detection procedure.

In the present study, we present an approach for detecting the buildings from high resolution pan-sharpened IKONOS imagery through binary SVM classification. In addition to original spectral bands of the IKONOS pan-sharpened image (Blue, Green, Red, NIR), the bands nDSM, NDVI, PC1, PC2, PC3, and PC4 are also included in the classification process. To

* Corresponding author.

assess the effect of sample size on the classification accuracy, the SVM classification was carried out using three different sample sizes of 500, 1000 and 2000.

2. THE METHODOLOGY

The flowchart for the proposed building detection procedure is given in Figure 1. First, a Digital Terrain Model (DTM) is generated from the existing vector data that contains contour lines and 3D points. Then, a DSM is generated from the stereo IKONOS Panchromatic image pairs. After generating the DTM and DSM, an nDSM is calculated by subtracting DTM from DSM. Next, the orthorectification of the pan-sharpened IKONOS image is carried out using the DSM. After that, the bands NDVI, PC1, PC2, PC3, and PC4 are calculated from the pan-sharpened IKONOS image. Finally, to detect the buildings, the orthorectified pan-sharpened IKONOS image along with the additional bands nDSM, NDVI, PC1, PC2, PC3, and PC4 is classified using the SVM classifier. After completing the classification, the artefacts are removed using the morphological operations and the non-building areas are masked out.

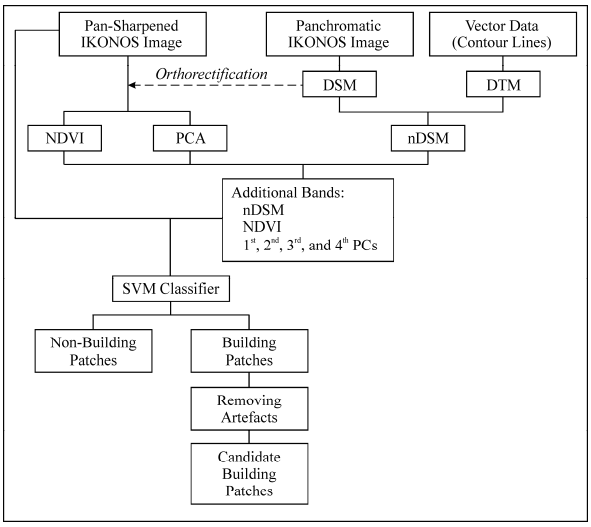


Figure 1. The flowchart of the building detection procedure.

3. THE STUDY AREA AND DATA SETS

The proposed building detection procedure was implemented in three selected areas in the Batikent district of Ankara, the capital city of Turkey (Figure 2). The areas show different residential and/or industrial characteristics. The first study area (Sub-Area I) covers the residential part of Batikent and includes buildings with different shapes, sizes, and dwelling types. The buildings falling within this area have different coloured roofs. The second study area (Sub-Area II) also represents the residential characteristics but compared to Sub-Area I it contains buildings with more regular shapes. The third study area (Sub-Area III) covers the industrial parts, where the roofs are usually in gray tone and compared to buildings in Sub-Area I and Sub-Area II the size of the buildings are larger.

The data sets used include the stereo panchromatic and the pan-sharpened IKONOS images and the existing building database, which was generated in 1999 from stereo aerial images. The IKONOS images were acquired on August 4, 2002. They were in *Geo* data format, which is the most economical product within the all IKONOS products. To evaluate the results, a reference building database was generated by updating the existing building database. To do that the missing and new buildings were visually detected on the screen and they were manually edited.

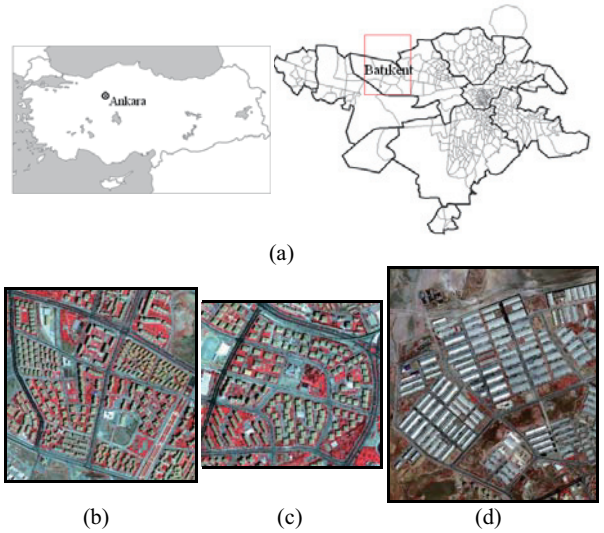


Figure 2. (a) The study area located in the Batikent district of Ankara. The Sub-Area I (b), Sub-Area II (c), and Sub-Area III (d).

4. GENERATING THE ADDITIONAL BANDS

In the present case, the bands nDSM, NDVI, PC1, PC2, PC3, and PC4 were also included in the classification. The band nDSM was included in the classification process so that it could help separate those man-made objects with different heights above the terrain. Similarly, the band NDVI was used as an additional band because it helps differentiate those buildings that are surrounded by trees and green vegetated areas. Our previous experiences have revealed that the industrial buildings with white or blue concrete roofs, which are dominant in Sub-Area III, are more easily detected on the band PC2, while the residential buildings with tile-roofs are more clearly differentiated on the band PC3. Therefore, it is evident that the inclusion of the Principal Components (PCs) in the classification process will increase the separation of the buildings from the non-building classes.

4.1 Calculating the Normalized Digital Surface Model

In the building detection procedure, the basic idea of using an nDSM is that the man-made objects with different heights above the terrain can be detected by applying a threshold to nDSM. A DTM is the elevation model of the landscape that does not include above ground objects. On the other hand, a DSM includes the objects with their heights above the ground as well as the topography. To calculate nDSM, a DSM was generated from the stereo panchromatic IKONOS images and a DTM was generated from the contour lines. Then, an nDSM was calculated by subtracting the DTM from the DSM. Next, to separate the above ground features from the terrain surface, a threshold value of 3m was applied to nDSM.

4.2 Orthorectification

To remove geometric distortions including relief displacement, both the panchromatic and pan-sharpened IKONOS satellite images were orthorectified. The orthorectification of the images was necessary so that the building areas to be extracted through SVM classification could be overlaid with the existing GIS database. The DSM used for the orthorectification procedure was the previously generated DSM from the stereo IKONOS panchromatic images.

4.3 Calculating the Normalized Difference Vegetation Index (NDVI)

The NDVI was used to detect the vegetated areas in the image. It was calculated using the red and near-infrared bands of the orthorectified IKONOS pan-sharpened (PSM) images (Eq.1);

$$NDVI = \frac{PSM4 - PSM3}{PSM4 + PSM3} \quad \text{(Eq. 1)}$$

where; *PSM 4* is the NIR band and *PSM 3* is the red band of the IKONOS pan-sharpened satellite image.

4.4 Principal Components Analysis

The Principal Component Analysis (PCA) is a technique designed to reduce redundancy in multispectral data (Lillesand and Kiefer, 1999). It is mainly used to reduce the number of bands. However, Singh and Harrison (1985) claim that the PCs may become more interpretable than the original data. Therefore, to include in the classification procedure the Principal Components were generated from the IKONOS pan-sharpened image bands. The spectral statistics of the IKONOS pan-sharpened image is given in table 1.

PC	Eigen Value	Std. Dev.	Variance (%)
1	132476.10	363.89	91.45
2	10295.38	101.46	7.11
3	1942.12	44.08	1.34
4	151.12	12.28	0.10

Table 1. The spectral statistics of the principal components.

5. SUPPORT VECTOR MACHINES (SVM) CLASSIFICATION

SVM is a supervised classification technique derived from statistical learning theory. The fundamentals of SVM were developed by Vapnik (1995). The SVM classification contains mainly three steps. First, the trainings of the classes are represented as feature vectors. Next, to perform the separation, these feature vectors are mapped into a feature space by using the kernel function. Finally, an n-dimensional hyperplane that optimally separates the classes are created.

To classify the pixels accurately, SVM develops a model using the training data of two separable classes with *i* samples represented by $(x_1, r_1), \ldots, (x_i, r_i)$, where $x_i \in R^n$ and $r \in \{1,-1\}$ in the n dimensional space. Here, x_i is the spectral value of the training data while r_i is the class label for a training case.

During the learning process, SVM optimizes the hyperplane position to have a maximum margin between the classes that are on different sides of the hyperplane. There may be many hyperplanes that separate the classes; however the aim is to find the optimum one and to maximize the margin. A hyperplane that separates the classes can be defined by the following equation (Eq. 2):

$$w.x + b = 0 \quad \text{(Eq. 2)}$$

where, w is the normal to the hyperplane, x is a point lying on the hyperplane, and b is the bias.

For the linearly separable two class case, a hyperplane can be defined as

$$w.x_i + b \geq +1 \quad for\ all\ r_i = +1, \quad \text{(Eq. 3)}$$

$$w.x_i + b \leq +1 \quad for\ all\ r_i = -1. \quad \text{(Eq. 4)}$$

The equations 3 and 4 can be combined into a single equation (Eq. 5);

$$r_i\,(w.x_i + b) - 1 \geq 0 \quad \text{(Eq. 5)}$$

The points on these hyperplanes are called the *support vectors*. The hyperplanes are used to define the optimal separating hyperplane. The optimal separating hyperplane is parallel to the hyperplanes and stays in the middle of them. The margin between these planes is $2/||w||$ ($||w||$ is the Euclidean norm of w). The maximization of this margin causes the optimization problem given below and equation 5 is the constraint:

$$min\ \{1/2\ ||w||2\} \quad \text{(Eq. 6)}$$

This problem can be defined by Lagrange multipliers:

Maximize:

$$\sum_{i=1}^{k} \alpha_i - \frac{1}{2} \sum_{i=1}^{k} \sum_{i=1}^{k} \alpha_i\, \alpha_i\, r_i\, r_i \left(x_i . x_j\right) \quad \text{(Eq. 7)}$$

Subject to:

$$\sum_{i=1}^{k} \alpha_i\, r_i = 0 \qquad \text{and } \alpha i \geq 0, \quad i = 1,2, \ldots,k \quad \text{(Eq. 8)}$$

The decision rule is then applied for classifying the data into two classes; +1 indicating one class and -1 the other class.

$$f(x) = sign\left(\sum_{sv} r_i\, \alpha_i^o (x_i . x) + b^o\right) \quad \text{(Eq. 9)}$$

If the classes in the dataset are mixed, the data may not be separated linearly. In the non-linear separable case, equation 7 cannot be fulfilled as constraint and therefore, slack variable (ξ) and penalty parameter (C) are defined. With the addition of these parameters the optimization problem and the constraints become:

$$min = \left[\frac{\|w\|^2}{2} + C\sum_{i=1}^{r} \xi_i\right] \quad \text{(Eq. 10)}$$

$$r_i\,(w.x_i + b) \geq 1 - \xi_i \quad \text{(Eq. 11)}$$

The penalty parameter (C) value is a form of regularization parameter and defines the trade off between the number of noisy training samples and classifier complexity. The approach can be adapted to allow for non-linear surfaces by transferring the training data into a high dimensional (feature) space (Figure 3).

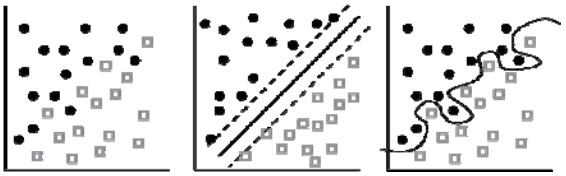
This time the problem can be defined by Lagrange multipliers using kernel function (K):

Maximize:

$$\sum_{i=1}^{k}\alpha_i - \frac{1}{2}\sum_{i=1}^{k}\sum_{i=1}^{k}\alpha_i\,\alpha_i\,r_i\,r_i\,K\left(x_i\,.\,x_j\right) \qquad \text{(Eq. 12)}$$

Subject to:

$$\sum_{i=1}^{k}\alpha_i\,r_i = 0 \quad \text{and} \quad 0 \le \alpha_i \le C,\ i = 1,2,\ldots,k \qquad \text{(Eq. 13)}$$



(a) (b) (c)

Figure 3. The processes of SVM classification. (a) The feature vectors of two classes in input space, (b) the feature vectors of two classes separated by a hyperplane in feature space, and (c) the separated two classes in the input space.

Using kernel function leads to change in decision rule as given below:

$$f(x) = sign\left(\sum_{sv} r_i\,\alpha_i^o K(x_i\,.\,x) + b^o\right) \qquad \text{(Eq. 14)}$$

There are different types of kernel types for different applications. The most common kernel types are *Linear*, *Polynomial*, *Radial Basis Function (RBF)*, and *Sigmoid*;

Linear: $K(x_i,x_j) = x_i T\,x_j$
Polynomial: $K(x_i,x_j) = (\gamma\,x_i T x_j + r)\,d,\ \gamma > 0$
RBF: $K(x_i,x_j) = exp(-\gamma\,||x_i - x_j||2),\ \gamma > 0$
Sigmoid: $K(x_i,x_j) = tanh\,(\gamma\,x_i T x_j + r)$

where:
γ is the gamma term in the kernel function for all kernel types except linear, d is the polynomial degree term in the kernel function for the polynomial kernel, r is the bias term in the kernel function for the polynomial and sigmoid kernels.

In the present study, the IKONOS pan-sharpened image covering the study areas was separated into the building and non-building classes through SVM classification. To perform the classification, the original and the derived bands of the IKONOS pan-sharpened image were labelled as follows:

B_1: Band Blue
B_2: Band Green
B_3: Band Red
B_4: Band NIR
B_{nDSM}: Band nDSM
B_{NDVI}: Band NDVI
B_{PC1}: Band PC1 (1st Principal Component)
B_{PC2}: Band PC2 (2nd Principal Component)
B_{PC3}: Band PC3 (3rd Principal Component)
B_{PC4}: Band PC4 (4th Principal Component)
By using these bands, fourteen image data sets were generated to perform the SVM classification (Table 2).

Data-Sets	Band Combinations
Data-Set 1	B_1, B_2, B_3, B_4
Data-Set 2	B_1, B_2, B_3, B_4, B_{nDSM}
Data-Set 3	B_1, B_2, B_3, B_4, B_{NDVI}
Data-Set 4	B_1, B_2, B_3, B_4, B_{PC1}
Data-Set 5	B_1, B_2, B_3, B_4, B_{PC2}
Data-Set 6	B_1, B_2, B_3, B_4, B_{PC3}
Data-Set 7	B_1, B_2, B_3, B_4, B_{PC4}
Data-Set 8	B_1, B_2, B_3, B_4, B_{nDSM}, B_{NDVI}
Data-Set 9	B_1, B_2, B_3, B_4, B_{nDSM}, B_{PC2}
Data-Set 10	B_1, B_2, B_3, B_4, B_{nDSM}, B_{PC3}
Data-Set 11	B_1, B_2, B_3, B_4, B_{nDSM}, B_{NDVI}, B_{PC2}
Data-Set 12	B_1, B_2, B_3, B_4, B_{nDSM}, B_{NDVI}, B_{PC3}
Data-Set 13	B_1, B_2, B_3, B_4, B_{PC1}, B_{PC2}, B_{PC3}, B_{PC4}
Data-Set 14	B_1, B_2, B_3, B_4, B_{nDSM}, B_{NDVI}, B_{PC1}, B_{PC2}, B_{PC3}, B_{PC4}

Table 2. The data sets and their band combinations used in the SVM classification.

After generating the image data sets, the training samples were collected from the representative homogeneous areas. In the present case, the class to be extracted was the building class, while the other classes including vegetation, road, bare land, shadow, and pavement composed the non-building class. An equal number of training pixels was collected for both the building and non-building classes. For the non-building class, the percentages of the training pixels were determined in accordance to the distribution of the class types. Initially, the training samples were collected manually. Then, 500, 1000, and 2000 pixels were randomly selected from the collected training areas. To make the assessments of the classified images, the test pixels were also collected from different locations than the training pixels as they must represent the unbiased reference information. We collected 2000, 4000, and 8000 test pixels, respectively for assessing the classifications conducted using 500, 1000, and 2000 training pixels.

For performing the SVM classification, the selection of the kernel method, determination of the C parameter, and the parameters related to the kernel are important. In the present case, the *Radial Basis Function (RBF)* was selected as the kernel method. This function can handle linearly non-separable problems and works well in most cases (ENVI Manual, 2006). In addition, γ was determined as the inverse of the number of bands in the input image and 1000 was taken for the value of the parameter C.

After performing the one-class SVM classification, the classified images contained two classes, one representing the building areas and the other representing the non-building areas. Next, the non-building areas were masked out from the classified images and therefore, the classified images contained only the building areas. Due to misclassification, the classified building areas contained artefacts. Therefore, these artefacts were removed using the morphological operations, which are image processing operations based on the shapes. In the present case, to eliminate those image details smaller than the structuring element, the *opening* and *closing* operations, which

are based on erosion and dilation operations, respectively, were used. Therefore, the global shape of the objects was not distorted (Sonka et al, 1998). An opening filter removes thin protrusions, outward pointing boundary irregularities, thin joins, and small isolated objects. On the other hand, a closing filter removes the thin gulf, the inward-pointing boundary irregularities and small holes (Gonzales et al, 2004). Therefore, combining a closing and an opening can be quite effective for removing the artefacts or noise.

6. THE RESULTS

The SVM classifier produced quite accurate results for the proposed building detection procedure. For all test areas, the overall accuracies were computed to be between 90% and 99%. Similarly, the Kappa Coefficients were found to be between 0.80 and 0.98. For Sub-Area I, Sub-Area II, and Sub-Area III, the assessment results of the SVM classification conducted using 500, 1000, and 2000 training samples are given in tables 3, 4, and 5 respectively. In figure 4, the classified images carried out using Data-Set 8 and 500 training samples are illustrated.

Sub-Area I	Overall Accuracy (%)			Kappa Coefficient		
	Train 500/ Test 2000	Train 1000/ Test 4000	Train 2000/ Test 8000	Train 500/ Test 2000	Train 1000/ Test 4000	Train 2000/ Test 8000
Data-Set 1	95.55	96.03	95.35	0.9110	0.9205	0.9070
Data-Set 2	98.70	98.48	98.65	0.9740	0.9695	0.9730
Data-Set 3	95.45	95.78	95.69	0.9090	0.9155	0.9137
Data-Set 4	95.25	95.75	95.31	0.9050	0.9150	0.9063
Data-Set 5	95.50	96.13	95.63	0.9100	0.9225	0.9125
Data-Set 6	95.80	96.05	95.60	0.9160	0.9210	0.9120
Data-Set 7	95.75	95.83	95.54	0.9150	0.9165	0.9107
Data-Set 8	98.60	98.43	98.60	0.9720	0.9685	0.9720
Data-Set 9	98.65	98.40	98.66	0.9730	0.9680	0.9732
Data-Set 10	98.85	98.73	98.64	0.9770	0.9745	0.9728
Data-Set 11	98.65	98.43	98.59	0.9730	0.9685	0.9718
Data-Set 12	98.70	98.73	98.68	0.9740	0.9745	0.9735
Data-Set 13	96.50	96.03	95.94	0.9300	0.9205	0.9187
Data-Set 14	98.65	98.35	98.71	0.9730	0.9670	0.9742

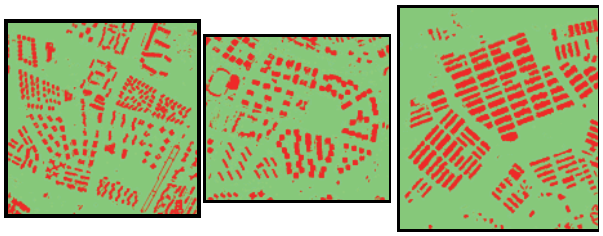
Table 3. The assessment results for Sub-Area I.

Sub-Area II	Overall Accuracy (%)			Kappa Coefficient		
	Train 500/ Test 2000	Train 1000/ Test 4000	Train 2000/ Test 8000	Train 500/ Test 2000	Train 1000/ Test 4000	Train 2000/ Test 8000
Data-Set 1	93.65	93.35	92.15	0.8730	0.8670	0.8430
Data-Set 2	96.50	97.15	97.16	0.9300	0.9430	0.9433
Data-Set 3	93.65	93.85	92.16	0.8730	0.8770	0.8433
Data-Set 4	93.65	94.10	91.84	0.8730	0.8820	0.8367
Data-Set 5	93.65	94.78	95.14	0.8730	0.8955	0.9028
Data-Set 6	93.80	94.83	95.50	0.8760	0.8965	0.9100
Data-Set 7	93.85	91.60	92.31	0.8770	0.8320	0.8462
Data-Set 8	96.05	96.88	96.98	0.9210	0.9375	0.9395
Data-Set 9	96.20	97.15	97.05	0.9240	0.9430	0.9410
Data-Set 10	96.30	96.70	97.15	0.9260	0.9340	0.9430
Data-Set 11	96.20	96.83	96.93	0.9240	0.9365	0.9385
Data-Set 12	96.20	96.68	96.59	0.9240	0.9335	0.9317
Data-Set 13	94.85	94.18	96.01	0.8970	0.8835	0.9203
Data-Set 14	94.50	94.90	94.88	0.8900	0.8980	0.8975

Table 4. The assessment results for Sub-Area II.

Sub-area III	Overall Accuracy (%)			Kappa Coefficient		
	Train 500/ Test 2000	Train 1000/ Test 4000	Train 2000/ Test 8000	Train 500/ Test 2000	Train 1000/ Test 4000	Train 2000/ Test 8000
Data-Set 1	90.05	91.95	92.64	0.8010	0.8390	0.8528
Data-Set 2	96.55	97.30	97.40	0.9310	0.9460	0.9480
Data-Set 3	90.10	91.63	92.04	0.8020	0.8325	0.8408
Data-Set 4	89.70	91.65	92.01	0.7940	0.8330	0.8403
Data-Set 5	92.90	93.28	93.46	0.8580	0.8655	0.8692
Data-Set 6	93.40	94.23	94.06	0.8680	0.8845	0.8812
Data-Set 7	90.70	91.98	92.83	0.8140	0.8395	0.8565
Data-Set 8	96.85	97.40	97.48	0.9370	0.9480	0.9495
Data-Set 9	95.20	97.28	97.31	0.9040	0.9455	0.9463
Data-Set 10	95.85	96.20	96.64	0.9170	0.9240	0.9327
Data-Set 11	96.65	97.88	97.85	0.9130	0.9575	0.9570
Data-Set 12	96.05	96.45	97.53	0.9210	0.9290	0.9505
Data-Set 13	94.70	95.10	95.15	0.8940	0.9020	0.9030
Data-Set 14	96.30	96.95	97.70	0.9260	0.9390	0.9540

Table 5. The assessment results for Sub-Area III.



(a) (b) (c)

Figure 4. The classified images for (a) Sub-Area I, (b) Sub-Area II (b), and (c) Sub-Area III.

For Sub-Area I, with the use of 500 training samples, the overall accuracies were computed in the range of 95.25% - 98.85%, while the Kappa Coefficients were computed in the range of 0.905 - 0.977. For Data-Sets 2, 8, 9, 10, 11, 12, and 14, the overall accuracies stayed above 98 %, while the other data sets provided relatively lower accuracies. With the use of 1000 training samples, the overall accuracies were computed in the range of 95.75% - 98.73%, while the Kappa Coefficients were computed in the range of 0.915 - 0.974. Similar to the results of 500-sample classification, the Data-Sets 2, 8, 9, 10, 11, 12, and 14 provided the best results with the overall accuracies higher than 98%. The SVM classification conducted using 2000 training samples provided the overall accuracies in the range of 95.31% - 98.71%, while the Kappa Coefficients were computed in the range of 0.906 - 0.974. In this case, the Data-Sets 2, 8, 9, 10, 11, 12, and 14 provided the best results with the overall accuracies higher than 98%.

For Sub-Area II, with the use of 500 training samples, the overall accuracies were computed in the range of 93.65% - 96.50%, while the Kappa Coefficients were computed in the range of 0.873 - 0.930. For Data-Sets 2, 8, 9, 10, 11, and 12, the overall accuracies were computed on the order of 96%, while the remaining data sets provided slightly lower values. With the use of 1000 training samples, the overall accuracies were computed in the range of 91.60% - 97.15%, while the Kappa Coefficients were stayed in the range of 0.832 - 0.943. In this case, the Data-Sets 2, 8, 9, 10, 11, and 12 provided the highest overall accuracies of above 96%, while the remaining data sets provided slightly lower values. The SVM classification conducted using 2000 training samples provided the overall accuracies in the range of 92.15% - 97.16%, while the Kappa Coefficients were computed in the range of 0.837 - 0.943. In this case, the Data-Sets 2, 8, 9, 10, 11, 12, and 13 provided the best results with the overall accuracies above 96%, while for the remaining data sets, the accuracies computed were slightly lower. The Data-Set4 provided the lowest accuracy of 91.84%.

For Sub-Area III, with the use of 500 training samples, the overall accuracies were computed in the range of 89.70% - 96.85%, while the Kappa Coefficients were computed in the range of 0.794 - 0.937. For Data-Sets 2, 8, 11, 12, and 14, the overall accuracies were computed on the order of 96%, while the remaining data sets provided slightly lower values. With the use of 1000 training samples, the overall accuracies were computed in the range of 91.63% - 97.88%, while the Kappa Coefficients were computed in the range of 0.832 - 0.948. In this case, Data-Sets 2, 8, 9, 11, and 14 provided the highest overall accuracies on the order of 97%, while for the remaining data sets, the computed accuracies were slightly lower. The Data-Set 3 provided the lowest accuracy of 91.63%. The SVM classification conducted using 2000 training samples provided

the overall accuracies in the range of 92.01% - 97.85%, while the Kappa Coefficients were computed in the range of 0.840 - 0.957. The Data-Sets 2, 8, 9, 11, 12, and 14 provided the overall accuracies of above 97%, while for the remaining data sets, the accuracies were computed to be slightly lower. In this case, the Data-Set 4 provided the lowest accuracy of 92.01%.

7. THE CONCLUSIONS

We showed in the present study that the inclusion of the additional bands nDSM, NDVI, and PCs in SVM classification increased the classification accuracy for detecting the buildings. While the overall accuracy was increased up to 7%, the Kappa Coefficients were increased up to 0.13. In specific, the inclusion of the band nDSM in SVM classification increased the accuracy up to 6.5%. Unexpectedly however, the inclusion of the band NDVI did not show a high percentage of increase in the classification accuracy. Similarly, the inclusion of each of the PCs in the classification process did not lead to any significant difference. However, including the all PCs at once in the classification process improved the overall accuracies up to 4%.

Concerning the effect of the training sample size on the classification accuracy we can state that increasing the training sample size increases the accuracy. However, in the present case the increase was not observed to be more than 3%. This is due to the fact that in SVM classification the location of the training samples in the feature space is more important than the training sample size.

ACKNOWLEDGEMENTS:

This research was supported by the State Planning Organization (DPT) Grants: BAP-08-11-DPT2002K120510.

REFERENCES:

Bellman, C. J., and Shortis, M. R., 2000. Building recognition using wavelet analysis and support vector machines. *Proceedings of SSC2003*.

Boyd, D. S., Sanchez-Hernandez, C. and Foody, G. M., 2006. Mapping a Specific Class for Priority Habitats Monitoring from Satellite Sensor Data. *International Journal of Remote Sensing*, Vol.27, No. 13, 2631-2644.

ENVI Manual, (2006), Version 4.3.

Foody, G. M. and Mathur, A., 2004. A relative Evaluation of Multiclass Image Classification by Support Vector Machines. *IEEE Transactions on Geoscience and Remote Sensing*, Vol.42, No. 6, pp. 1335-1343.

Foody, G. M., Mathur, A., Sanchez-Hernandez, C. and Boyd, D. S., 2006. Training Set Size Requirements for the Classification of a Specific Class. *Remote Sensing of Environment*, 104, pp.1-14.

Gonzalez, R. C., Woods, R. E. and Eddins, S. L., 2004. *Digital Image Processing Using Matlab*. Pearson Education, Inc. Upper Saddle River, New Jersey.

Huang, C., Davis, L. S., and Townshend J. R. G., 2002. An Assessment of Support Vector Machines for Land Cover Classification. *International Journal of Remote Sensing*, 23(4), pp. 725–749.

Huang, X., Zhang, L., and Li, P., 2007. An adaptive multiscale information fusion for feature extraction and classification of IKONOS multispectral imagery over urban areas. *IEEE Geoscience and Remote Sensing Letters*, 4(4), pp.654-658.

Li, H., Gu, H., Han, Y., and Yang, J., 2010. Object-oriented classification of high-resolution remote sensing imagery based on an improved colour structure code and a support vector machine. *International Journal of Remote Sensing*, 31(6), pp. 1453–1470.

Lillesand, T.M. and Kiefer, R.W., 1999. *Remote Sensing and Image Interpretation*. 4th edition, John Wiley & Sons, Inc.

Pal, M. and Mather, P. M., 2005. Support Vector Machines for Classification in Remote Sensing. *International Journal of Remote Sensing*, 26(5), pp.1007–1011.

Sanchez-Hernandez, C., Boyd, D. S. and Foody G., M., 2007. One-Class Classification for a Specific Land-Cover Class: SVDD Classification of Fenland. *IEEE Transactions on Geoscience and Remote Sensing*, 45(4), pp.1061-1073.

Singh, A., and Harrison, A., 1985. Standardized Principal Components. *International Journal of Remote Sensing*, 6, pp. 883-896.

Sonka, M., Hlavac, V. and Boyle, R., 1998. *Image Processing, Analysis, and Machine Vision.* PWS Publishing

Tuia, D., Ratle, F., Pozdnoukhov, A., and Camps-Valls, G., 2010. Multisource composite kernels for urban-image classification. IEEE Geoscience and Remote Sensing Letters, 7(1), pp.88-92.

Vapnik, V. N., 1995. *The Nature of Statistical Learning Theory*. New York: Springer-Verlag, ISBN 0-387-94559-8.

Watanachaturaporn, P., Arora M. K., Varshney P. K., 2008. Multisource Classification Using Support Vector Machines: An Emprical Comparison with Decision Tree and Neural Network Classifiers. *Photogrammetric Engineering and Remote Sensing*, 74(2), pp. 239 – 246.

Zhu, G. and D. G. Blumberg, 2002. Classification Using ASTER Data and SVM Algorithms: The case study of Beer Sheva, Israel. *Remote Sensing of Environment*, 80(2), pp. 233–240.