IMPERVIOUS SURFACE AREAS CLASSIFICATION FROM GEOEYE-1 AND WORLDVIEW-2 SATELLITE IMAGERY THROUGH OBIA APPROACH IN A COASTAL AREA OF ALMERIA (SPAIN)

Ismael Fernández-Luque[1], Fernando Aguilar-Torres[1], Manuel Aguilar-Torres[1], Flor Álvarez-Taboada[2] and Mar Saldaña-Díaz[1]

1. University of Almería, Department of Engineering, Almería, Spain; ismaelf@ual.es, faguilar@ual.es, maguilar@ual.es, marsalda@hotmail.com
2. University of León, Department of Mining Engineering, Ponferrada, Spain; flor.alvarez@unileon.es

Abstract

GeoEye-1 and WorldView-2 satellite imagery were used to classify impervious surface areas (ISAs) in a coastal area of Almeria (south-east of Spain). ISAs can play a key role in ecosystem health as well as in loss of sediment discharge and shoreline recession. An object-based image analysis (OBIA) approach and support vector machines (SVM) classifier were used. The feature vectors utilized included basic band information, ratios between bands and texture information.

The Kappa test was performed to check the statistically significant differences between all the feature sets for each satellite imagery and a separability matrix was constructed. Thus, the obtained accuracy results were compared in order to test which satellite imagery achieved the highest accuracy. Similarly, those results were compared with a similar experiment carried out on archival aerial orthoimages which covered the same area.

Finally, the sampling method was tested. First, a pilot area was used to classify two different zones. Secondly, an ad hoc training was employed for each separated zone. Accuracy classification results proved that the distribution and number of samples were relevant for the final accuracy.

INTRODUCTION

From the second half of the twentieth century the urbanization processes have been affecting a large coastal area in Spain, mainly in the Mediterranean coast. The increase of the population as well as the touristic activities that take place in those areas have played a key role for such development. Impervious surface areas (ISAs) are defined as anthropogenic features thorough water cannot infiltrate into the soil (1) and they are a good indicator of the degree of urbanization in one area. ISAs influence the hydrology of the watershed, impacting on the runoff features and increasing the stormflow (2). Thus, determination of ISA percentage is crucial for the evaluation of non-point runoff and the estimate of the water quality. Additionally, the percentage of ISA is correlated with the health of the ecosystem (3) which could be classified as stressed (up to 10%), impacted (from 11 to 25%), and degraded (more than 25%).

Mainly due to the development of the very high spatial resolution (VHR) satellite imagery, pervious/impervious surfaces identification can be undertaken by means of Remote sensing techniques and image classification approaches (4). Platforms such as IKONOS, QuickBird, GeoEye-1 or WorldView-2 allow the application of classification approaches over high spatial resolution images, with a ground sample distance from 1 meter and even better in their panchromatic bands (PAN). Moreover, those satellite images include some spectral bands, such as Near-Infrared (Nir), which can improve classification significantly as compared to archival aerial orthoimages, which are tipically composed by red, green, and blue (RGB) bands. That novelty makes important to check the capability of the new satellite imagery for ISAs classification. In this sense, this work tests an object-based image analysis (OBIA) by applying a non-parametric

classification approach on GeoEye-1 and WorldView-2 orthoimages corresponding to a coastal area of Almeria, Spain.

The widely known “salt and pepper” effect (caused by per-pixel classification approaches mainly in VHR image with a high local variance) can be surpassed by OBIA approaches (5). Since urban environments consist of small features as well as very different materials, OBIA can lead to an improvement of the performance of supervised classifiers when a relatively small ground sample distance (GSD) is used (6). OBIA is based on creating meaningful objects by means of aggregating individual pixels according to some parameters, so the number of units to be classified is significantly reduced. Furthermore, topological and contextual based features can be used for classification purposes.

Regarding the classification methods, non-parametric classifiers are preferred for many applications since they avoid the previous assumptions about data distribution. In that sense, Support Vector Machines (SVM) approach, which tries to find a hyperplane that splits a training data set into two subsets by using a minimum number of samples (7), was tested in this work to classify the scene.

A key step for image classification is the determination of the most appropriate feature set to use. Here, basic spectral information, indices between bands, and basic texture information were tested. Then, the most suitable feature set was chosen in order to robustly and statistically differentiate the feature sets used by means of the Kappa test.

The approach described here is based on a general methodology previously developed on an archival aerial orthoimage of the same area (8) and, therefore, the accuracy of both classifications, GeoEye-1 image and archival orthoimages, could be compared.

MATERIALS AND METHODS

The study site of this work was placed on a heavily urbanized coastal area in the province of Almeria, between the villages of Villaricos and Garrucha, located at the South-East of the Spanish Mediterranean coast. This area comprised approximately a coastal fringe of 11 km long and 775 m wide. For training strategies reason, the whole area was divided in three parts: one pilot area located in the North, which was used for feature set selection purposes, and the areas A and B dividing the remaining study site in two additional parts (Figure 1).

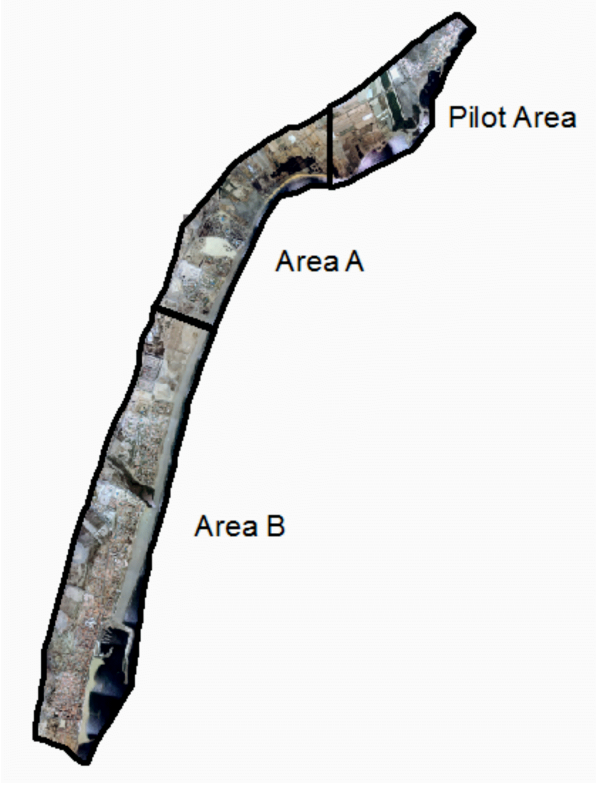

Figure 1: Distribution of the different experimental areas.

GeoEye-1 data

This satellite imagery is comprised of the panchromatic image (PAN) and also of the bands Red, Green, Blue, and Near Infrared (R, G, B, Nir) from the multispectral image (MS). The original ground sample distance (GSD) was 0.41 m and 1.65 m for PAN and MS images respectively. However, those GSD were down-sampled to 0.5 m and 2 m due to US governmental requirement. An image taken with an off-nadir of 9º was used for this work and two different orthoimages were obtained from the original Geo images: first, one PAN orthoimage was derived by means of a zero order rational polynomial coefficients sensor model (RPC0) refined with 7 accurate ground control points (GCPs) and a high accurate LiDAR-derived DEM; and second, a MS orthoimage was obtained based on a pan-sharpened image with 0.5 m GSD and containing the spectral information gathered from the MS image. Thus, the geopositioning accuracy, evaluated as $RMSE_{2D}$ computed on 48 independent check points (ICPs), turned out to take a value of 0.41 m.

WorldView-2 data

Likewise GeoEye-1 (GE1), WorldView-2 (WV2) provides PAN and MS images. However, MS image adds for additional bands to those provided by GeoEye-1: Coastal Blue, Yellow, Red Edge, and Near Infrared-2 (CB, Y, RE, Nir2). The final GSD is the same than that for GE1 (also for the same governmental reasons) but the original one varies, being 0.46 m and 1.84 m for PAN and MS images respectively. WV2 images were acquired in August 2011, that is only 9 days later the GE1 ones, presenting an off-nadir angle of 10º. Following the same scheme than GE1 (i.e. RPC0 supported by 7 GCPs and using same ancillary LiDAR-derived DEM), WV2 orthoimage was achieved with a $RMSE_{2D}$ of 0.46 m (evaluated on the same ICPs set).

For further and detailed information about the orientation and orthorectification processes for both satellite images the reader is referred to the previous works recently published by our research group (9,10)

Object-based approach

As previously mentioned, the minimum classification unit used for this work was the object. Thus, every object in the scene was classified as pervious or impervious according with the features estimated by all pixels within the object and every training and validation sample corresponded to an only object. Segmentation process was carried out through the multiresolution segmentation algorithm running under eCognition 8® software (11) using all bands (equally-weighted) included in the pan-sharpened and PAN orthoimages. The segmentation method parameters shape and compactness for both classifications were 0.3 and 0.7, respectively. However, the scale parameter (which controls the size of the objects based on the previous parameters and the bands information used) was 50 for GE1 and 40 for WV2 in order to obtain a similar number of total objects that assured similar meaning and size of the objects.

Classification and sampling strategies

According to the results of (8), a direct classification strategy was followed which means that several classes (or sub-classes) were grouped in the target super-classes pervious and impervious depending on their land use. For instance, samples from the classes such as buildings or roads were gathered into impervious samples, whereas sand-beaches or bare soil samples were grouped into pervious samples. Thus, a direct classification of the target classes could be achieved without taking into account their internal variability.

The sampling method was the randomly stratified for both training and validation samples. All the considered sub-classes were used for this sampling in order to keep a balanced representation in the scene. Once this method was applied for GE1 classification, it was tried to keep the same sample location for WV2 in order to carry out the same strategy for both satellite images because of comparison purposes. Finally, 215 training and 728 validation samples were collected for the pilot area, which seems to be a reasonable size according with (12).

Feature sets

Table 1 shows the feature sets used for this experiment. Basic information included the mean digital number (DN) of every pan-sharpened and PAN image band although two different basic feature sets were considered for WV2: first, only coincident bands (R, G, B, and Nir) were used ('Basic1'); and secondly, additional bands from WV2 were added to the previous one ('Basic2'). In the case of the 'Rates' feature set, chromatic ratios for RGB bands were considered in which every band was divided by the addition of those three bands -e.g. the red ratio is R/(R+G+B)-, while normalized differences of all bands and also the green-red ratio (G/R) were also used. All rates were added to the basic feature set. Furthermore, some texture features based on the local variance were added to the previous feature set by varying the local window size in order to check its influence on classification accuracy. In this sense, 3x3, 5x5, and 7x7 windows were utilized (represented in Table 1 by T3, T5, and T7, respectively). Finally, all the window sizes were added to the 'Rates' feature set (named 'TextureAll') and to the basic feature sets (named 'TextureOnly') in order to create two new feature sets.

Table 1: Feature sets used for GE1 and WV2 classifications. Figures in brackets indicate the number of features used.

Feature set name	GeoEye-1	WorldView-2
Basic1	PAN + R + G+ B + Nir (5)	PAN + R + G+ B + Nir (5)
Basic2	-	Basic1 + CB + Y + RE + Nir2 (9)
Rates	Basic1 +Chromatic ratios (R, G, B) + G/R ratio + NDBI + NDGI + NDVI (12)	Basic1 +Chromatic ratios (R, G, B) + G/R ratio + NDBI + NDGI + NDVI + NDYI + NDREI + NDCBI (19)
Texture3	Rates + T3 (13)	Rates + T3 (20)
Texture5	Rates + T5 (13)	Rates + T5 (20)

Texture7	Rates + T7 (13)	Rates + T7 (20)
TextureAll	Rates + T3 + T5 + T7 (15)	Rates + T3 + T5 + T7 (22)
TextureOnly	Basic1 + T3 + T5 + T7 (8)	Basic2 + T3 + T5 + T7 (12)

Classification algorithm

Support vector machines (SVMs) approach was reported as an appropriate non-parametric classifier for ISAs classification in a previous study (8). SVMs achieved more accurate results than classification and regression trees (CART) and it was proved more efficient than nearest neighbour (NN) classifier although NN yielded similar classification accuracy results. Therefore, only SVM was considered in this work. Summing up, SVM is based on finding a hyperplane which splits a data set (a feature set in this case) into two subsets during the training stage. That training stage obtains an optimum boundary decision solution that minimizes misclassifications. However, only a few number of training samples are used to define the hyperplane (called support vectors), so a smaller number of samples than other classifiers are expected to be required. The algorithm needs a kernel function. Therefore, the widely known radial basic function (RBF) was utilized in this work by using the free-distribution library LIBSVM (13).

Accuracy and significant difference assessment

Overall accuracy (OA), producer's accuracy (PA), user's accuracy (UA), and KHAT statistic (and its variance) were estimated from the error matrices derived from the validation samples for all the classification projects carried out in the pilot area. Moreover, Kappa test based on the KHAT comparison was performed in order to select the most appropriate feature set. According to (12), when the test statistic (namely Z) is larger than 1.96, the KHAT difference can be considered as statistically different at a 95% confidence level (normality is assumed). Please, see (12) to check all the formulation. Using the Kappa test, significant differences between the image sources (GE1 or WV2) can be also estimated.

Classification of External Areas A and B

The most suitable feature set derived from the pilot area results was used for the classification of the external areas A and B. However, two strategies were followed for the training phase. Firstly, the pilot area training samples were used for the external areas classification and, secondly, two *ad hoc* training samples were collected on each external area to an on-site classification. The same validation samples for both training samples were used for the accuracy assessment and comparison for each area. Therefore, Kappa test was able to be undertaken for pilot area and *ad hoc* training comparison. Finally, all training and tested samples from the three areas were grouped and the entire study site was classified.

RESULTS

GeoEye-1 feature set selection

General accuracy results for GE1 are presented in *Table 2.* The OA results indicated that all the feature sets applied exceeded the required minimum 85% OA suggested in (12). However, large differences between PA1 (Producer's accuracy for pervious class) and PA2 (Producer's accuracy for impervious class) can be observed, excepting the 'TextureAll' feature set which presented a more balanced PAs and UAs results. Regarding the differences through Kappa test, only 'TextureAll' was statistically more accurate than one other feature set ('Rates') while any other set was statistically different. Therefore, 'TextureAll' was chosen for GE1 classification.

Table 2: Accuracy assessment results for GE1 experiment in the pilot area. OA, PA, and UA are expressed in % and 1 and 2 indicate pervious and impervious class, respectively.

Feature set	OA	PA1	PA2	UA1	UA2	KHAT

Basic	87.8	95.9	73.7	86.4	91.2	0.725
Rates	86.3	95.5	70.3	84.8	89.9	0.689
TextureOnly	87.4	93.5	76.7	87.4	87.2	0.720
Texture3	88.6	96.5	74.8	86.9	92.6	0.744
Texture5	88.9	97.2	74.4	86.8	93.8	0.749
Texture7	87.5	94.6	75.2	86.9	88.9	0.721
TextureAll	90.0	93.7	83.5	90.8	88.4	0.781

WorldView-2 feature set selection

Table 3 shows the general accuracy results for WV2 classification. Again, PAs yielded quite different results for both target classes while UAs resulted in smaller differences. Although it was not the most accurate regarding the OA results, 'TextureAll' feature set was again chosen since no statistical differences could be found with respect to 'Texture7' or 'TextureOnly'. For WV2, more statistically significant differences were found and both 'Basic1' and 'Basic2' feature sets were less accurate than all texture feature sets except for 'Texture3' (*Table 4*).

Table 3: Accuracy assessment results for WV2 experiment in the pilot area. OA, PA, and UA are expressed in % and 1 and 2 indicate pervious and impervious class, respectively.

Feature set	OA	PA1	PA2	UA1	UA2	KHAT
Basic1	86.4	96.5	68.8	84.3	92.0	0.690
Basic2	87.4	95.5	73.3	86.1	90.3	0.716
Rates	88.0	89.6	85.3	91.4	82.5	0.744
TextureOnly	90.8	93.5	86.1	92.1	88.4	0.800
Texture3	89.0	94.2	80.1	89.1	88.8	0.758
Texture5	90.8	96.3	81.2	89.9	92.7	0.796
Texture7	91.5	96.1	83.5	91.0	92.5	0.812
TextureAll	90.8	95.0	83.5	90.9	90.6	0.798

Table 4.Separability matrix for WV2 feature sets. Kappa test results are depicted and bold figures highlight statistically different pair comparisons (Z>1.96; p<0.05)).

	Texture7	TextureOnly	TextureAll	Texture5	Texture3	Rates	Basic2	Basic1
Texture7	0							
TextureOnly	0.37	0						
TextureAll	0.44	0.07	0					
Texture5	0.50	0.13	0.06	0				
Texture3	1.60	1.23	1.16	1.10	0			
Rates	**1.99**	1.63	1.56	1.49	0.38	0		
Basic2	**2.72**	**2.36**	**2.28**	**2.22**	1.12	0.75	0	
Basic1	**3.38**	**3.02**	**2.95**	**2.88**	1.79	1.41	0.66	0

GE1 vs WV2 comparison

Separability results are shown in *Table 5*. The differences were only significant for 'Texture7' and 'TextureOnly' feature sets, WV2 yielding more accurate results. When classification accuracy differences were not significant, WV2 KHAT values were higher than those for GE1, excepting the basic feature sets. That could imply that normalized differences of WV2 extra bands included in the 'Rates' feature set can play an important role.

Table 4. Separability results between GE1 and WV2 classifications. Bold figures indicate significant differences (Z>1.96; p<0.05). 'Basic2' feature set for WV2 was compared with 'Basic1' for GE1.

	KHAT	Z statistic
GE1_Basic1	0.725	0.90
WV2_Basic1	0.690	
WV2_Basic2	0.716	0.24
GE1_Rates	0.689	1.43
WV2_Rates	0.744	
GE1_Texture3	0.744	0.39
WV2_Texture3	0.758	
GE1_Texture5	0.749	1.35
WV2_Texture5	0.796	
GE1_Texture7	0.721	**2.58**
WV2_Texture7	0.812	
GE1_TextureOnly	0.720	**2.25**
WV2_TextureOnly	0.800	
GE1_TextureAll	0.781	0.51
WV2_TextureAll	0.798	

It has been seen that the differences were not too high for the accuracy results. Additionally, the variable window size for computing the local variance texture indicator did not significantly affect the final classification accuracy results for any data source, finding that the combination of different window sizes could improve it. If results are compared with those reported in (8) from a high resolution RGB archival orthoimage, it is important to note that no differences could be established when including the texture feature (KHAT was 0.776 for that experiment).

External areas classification

'TextureAll' was the chosen feature set for both satellites according to the classification results in the pilot area. External areas (A and B) were classified using both pilot area and *ad hoc* training. The statistical differences are depicted in *Table 5*. A clear difference can be observed from this table. While for GE1 *ad hoc* results were significantly more accurate than those from pilot area training, the WV2 results showed that no differences could be detected, although KHAT results for *ad hoc* classification were higher. Therefore, WV2 data set can be considered more consistent since a local training set was able to accurately classify external areas although on-site training samples are still recommended. If all the results were merged, only the *ad hoc* KHAT on area B for GE1 could be considered more accurate than all the WV2 results. In the same way, the pilot area KHAT values for WV2 were significantly more accurate than the pilot area KHAT values for GE1.

Table 5. Separability matrices for areas A and B in GE1 and WV2 experiments. PA denotes pilot area training while AH denotes ad hoc training set.

GE1 experiment

		B_AH	A_AH	B_PA	A_PA
	KHAT	0.804	0.732	0.515	0.505
B_AH	0.804	0			
A_AH	0.732	1.84	0		
B_PA	0.515	**6.55**	**4.67**	0	

WV2 experiment

	B_AH	A_AH	B_PA	A_PA
KHAT	0.725	0.713	0.689	0.670
0.725	0			
0.713	0.28	0		
0.689	0.83	0.55	0	

A_PA	0.505	**6.66**	**4.81**	0.20	0	0.670	1.24	0.96	0.41	0

CONCLUSIONS

The importance of including illumination invariant features such as texture on land use classification has been proved in this work. However, low differences in classification accuracy were estimated for GeoEye-1 and WorldView-2 images if compared with the feature sets formed by the basic bands information. Generally, the results achieved for ISAs classification can be graded as accurate enough. For example, (14) achieved an OA about 92% when classifying ISAs by means of true colour aerial photography and OBIA approach and (15) attained an OA of 90% by applying SVM on a Quickbird image. Additionally, and according with (15), a limitation in the maximum accuracy can exist when OBIA approach is used since under-segmentation errors could occur. In this work, an accuracy of 90% seemed to be around the maximum value to obtain from the starting segmentation since additional features did not lead to a higher accuracy. Therefore, the accuracy results obtained in this work seems to be reasonably accurate as compared to the available literature. Furthermore, no large difference between both data source could be found when the same feature sets were compared for both images. Finally, the influence of the training sample set was proved since an external area can have a key influence as could be proved for GE1 study. However, WV2 achieved better results in case of classifying areas by an external training sample set. Summing up, an on-site sampling for training purposes is recommended regardless of the data source used.

ACKNOWLEDGEMENTS

This work takes part of the general research lines promoted by the CEI-MAR Campus of International Excellence as a joint initiative between the universities of Almeria, Granada, Huelva and Málaga, headed by the University of Cádiz (further information can be retrieved from www.campusdelmar.es). This work was supported by the Andalusia Regional Government, Spain, through the Excellence Research Project RNM-3575 and the Spanish Ministry for Science and Innovation (Spanish Government) under Grant Reference CTM2010-16573. Both projects are co-financed by the European Union through the European Regional Developed Funds (FEDER).

References

1 Weng Q, 2012. Remote Sensing of impervious surfaces in the urban areas: Requirements, methods, and trends. Remote Sensing of Environment, 15 (117): 34-49.

2 Gillies R R, J B Box, J Symanzik & E J Rodemaker, 2003. Effects or urbanization on the aquatic fauna of the Line Creek watershed, Atlanta-A satellite perspective. Remote Sensing of Environment, 3 (86): 411-422.

3 Arnold C L & C J Gibbons, 1996. Impervious surface coverage: The emergence of a key environmental indicator. Journal of American Planning Association, 2 (62): 243-258.

4 Lu D & Q Weng, 2009. Extraction of urban impervious surfaces from an IKONOS image. International Journal of Remote Sensing, 5 (30): 1297-1311.

5 Pu R, S Landry & Q Yu, 2011. Object-based urban detailed land cover classification with high spatial resolution IKONOS imagery. International Journal of Remote Sensing, 12 (32): 3285-3308.

6 Blaschke T, 2010. Object based image analysis for remote sensing. ISPRS Journal of Photogrammetry and Remote Sensing, 1 (65): 2-16.

7 Mountakis G, J Im and C Ogole, 2011. Support Vector Machines in remote sensing: A review. ISPRS Journal of Photogrammetry and Remote Sensin, 3 (66): 247-259.

8 Fernández I, F J Aguilar, M F Álvarez & M A Aguilar, *In press*. Non-parametric object-based approaches to carry out ISA classification from archival aerial orthoimages. IEEE Journal of selected topics in applied earth observations and remote sensing

9 Aguilar M A, F J Aguilar, M M Saldaña & I Fernández, 2012. Geopositioning accuracy assessment of GeoEye-1 panchromatic and multispectral imagery. Photogrammetric Engineering and Remote Sensing, 3 (78): 247-257.

10 Aguilar M A, M M Saldaña & F J Aguilar, 2013. Assessing geometric accuracy of the orthorectification process from GeoEye-1 and WorldView-2 panchromatic images. International Journal of Applied Earth Observation and Geoinformation , 1 (21): 427-435.

11 Benz U C, P Hofmann, G Willhauck, I Lingenfelder & M Heynen, 2004. Multi-resolution object-oriented fuzzy analysis of remote sensing data for GIS-ready information. ISPRS Journal of Photogrammetry and Remote Sensing, 3-4 (58): 239-258.

12 Congalton R G & K Green, 2009. Assessing the accuracy of Remotely Sensed Data (Boca Raton) 210.

13 Chang C & C Lin, 2011. LIBSVM: A library for support vector machines. ACM Transactions on Intelligent Systems and Technology, 2.

14 Blaschke T, S Lang & G Hay, 2008. Object-based image analysis: spatial concepts for knowledge-driven remote sensing, section 6.2 (Springer-Verlag) 817.

15 Liu D & Xia F, 2010. Assessing object-based classification: Advantages and limitations, Remote Sensing Letters, 4 (1): 187-194.