ROAD EXTRACTION METHOD AT THE PIXEL AND OBJECT LEVEL USING HIGH RESOLUTION IMAGES AND LIDAR DATA WITH EVIDENCE THEORY

Borja Rodríguez-Cuenca [(1)], Álex Martínez de Agirre [(1)], María Concepción Alonso [(1)], Alberto del Val[(1)]

[(1)] Alcalá University, Campus Universitario, Ctra. Madrid-Barcelona, Km. 33,600, 28805, Email:borja.rodriguezc@uah.es

ABSTRACT

Light Detection and Ranging (LIDAR) data provides accurate height information for objects on the Earth's surface. It is common to combine LIDAR data with satellite or aerial imagery to determine the location of different cartographic entities. In this work, an automated method for road extraction in urban areas from high resolution aerial images is presented. This method is based on the Dempster-Shafer Theory of evidence, which consists of fusing information from different information sources. The proposed method is applied to road extraction in two different ways: at the pixel level and at the object level. The results provided for both levels are compared with a ground truth created by the authors of this work in order to determine which method provides the best accuracy.

1. INTRODUCTION

Land Use Land Cover (LULC) automatic detection from remote sensing data has been established as an indispensable tool for providing adequate information for natural resource management and sustainable development decision makers. LULC changes are one of the main factors influencing the evolution of landscapes and it is important to detect these changes as automatically as possible. In the past, several methods have been developed to detect changes in the LULC of different landscapes. However, it is hard to find an appropriate method to detect LULC in every region, especially in urban areas where different types of confluence uses (residential, commercial, industrial, leisure) occur. To generate LULC databases, several programs at the national or international level are available, such as the CORINE Land Cover database in Europe or the Sistemas de Información sobre Ocupación del Suelo en España (SIOSE) database in Spain.

The main goal of LULC detection is to automate the detection process as much as possible. Several works based on thresholding or classification have tried to automate the LULC detection from remote sensing data. The thresholding method consists of using the spectral behavior of every LULC object in order to detect it from aerial or satellite images. Thresholding establishes a decision rule in every index that allows for determining the location of every LULC object that is considered. Some of these indices are the Normalized Difference Vegetation Index (NDVI) [1], the Soil Adjusted Vegetation Index (SAVI) [2], and the Burned Area Index (BAI) [3]. The most common way to apply the decision rules is by using a decision tree [4]. Some works that use the thresholding method are [5], where buildings are detected using a rule combining LIDAR data and aerial images, [6] that classifies LANDSAT images into five LULC categories using decision trees or [7], where the decision trees method is improved and is compared with a classification algorithm. Otherwise, image classification consists of dividing an image into different classes according to its spectral characteristics. While several algorithms can be used to perform a classification, the most widely used classification algorithms include Maximum Likelihood [8], Support Vector Machine (SVM) [9], and neural network algorithms [10]. Examples of some of the studies that have used algorithm classifications to detect LULC are: [11], which detected eight types of land cover using the SVM supervised classification; [12], which compares the performance of neural network and maximum likelihood; and [13], in which a supervised and an unsupervised classification are carried out.

In the current literature, there are works based on the extraction of different LULC. The present work is focused on the automatic detection of roads in urban regions from aerial or satellite images and LIDAR data. Several methods have been used to extract roads from aerial or satellite images [14] or with radar and LIDAR data [15]. A review of different methods to detect roads from remote sensing is presented in [16][17][18].

In the work presented in this paper, thresholding is carried out on four decision indexes in order to automatically identify every existing road in the study area using the Dempster-Shafer Theory. The method has been applied at the pixel level and at the object level, in order to compare the results obtained from both levels to determine which provides the best results.

2. MATERIAL

In this document, an aerial image and LIDAR data have

been used as inputs. The aerial image has a spatial resolution of 1 meter/pixel and is formed by four spectral bands: blue (B), green (G), red (R), and near infrared (NIR). The image was taken in summer 2010 in the Spanish city of Alcalá de Henares. The scene used in this paper has a resolution of 400x400 pixels (Fig.1).

LIDAR technology allows the calculation of the digital surface model (DSM) with precision in the order of 20 cm in planimetry and 30 cm in height [19]. The combination of a sweeping beam laser with inertial navigation systems and a GPS guarantees a high geometric precision in the data. In this study, we used the LEICA ALS50-II (Leica Geosystems AG, Heerbrugg, Switzerland) sensor to capture the LIDAR data. Although a DSM could also be obtained by using classical digital photogrammetry, that method would be more expensive and slower to produce. The LIDAR data used in the present work correspond to a flight, conducted in summer 2010. The flight was taken at an altitude of approximately 1800 m and with a minimum point density of 0.5 points per square meter. The coordinate system is WGS84 with orthometric heights.

Figure 1. Aerial image of the studied area in true color RGB

3. METHOD

In this paper, a method for the automatic detection of roads in urban areas is presented. Identification and extraction began with the reading of the aerial image and the rasterization of LIDAR data. After that, a segmentation of the aerial image into regions was carried out. The Dempster-Shafer theory of evidence was applied to determine which category corresponded to each region of the segmented image and to every pixel of the original image, depending on the value of every studied index. Fig. 2 shows a flow chart of the method proposed in this paper.

The first step in the procedure is the reading of the different bands involved in the process. From the aerial image, three optic bands—red (R), green (G), and blue (B)—and the near infrared band (NIR) are used.

Height and intensity information was obtained from the LIDAR data (Fig. 3). The Digital Surface Model (DSM) was obtained from the first echo of each pulse after filtering the noise that was generated in the process of data capture. The DSM contained information regarding construction, vegetation, and uncultivated ground. The digital terrain model (DTM) was generated as a product derived from the DSM after employing a semi-automatic method developed by [20]. Finally, calculation of the difference between the DSM and the DTM generated the standard DSM (nDSM, Normalized Digital Surface Model). Apart from height, intensity data was rasterized in order to add information to the process. As in the rasterization of the height data, the final intensity of each pixel was calculated as the average of all LIDAR points contained in each cell.

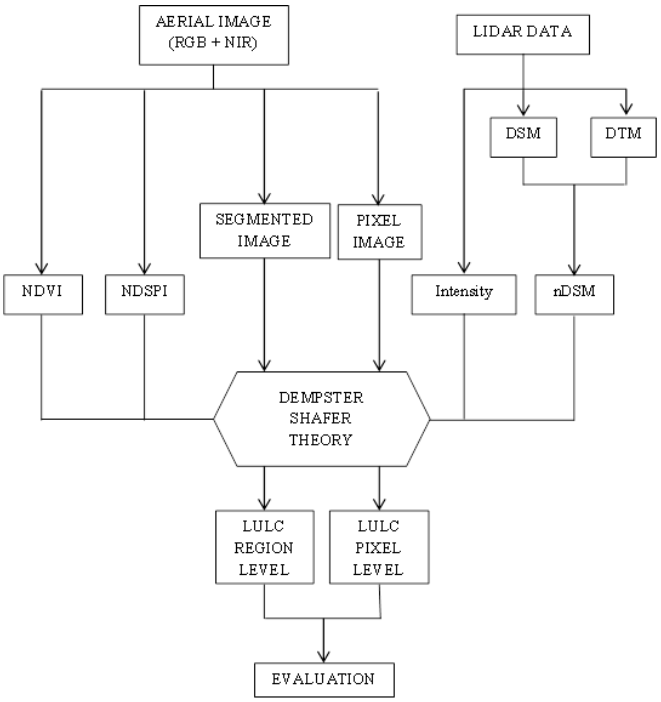
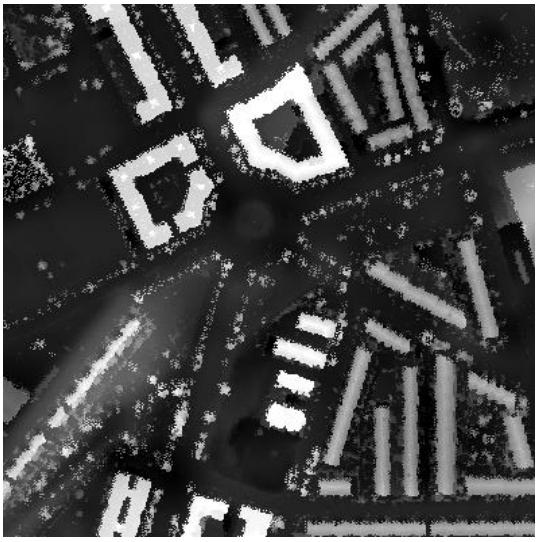


Figure 2. Flow chart of the proposal algorithm

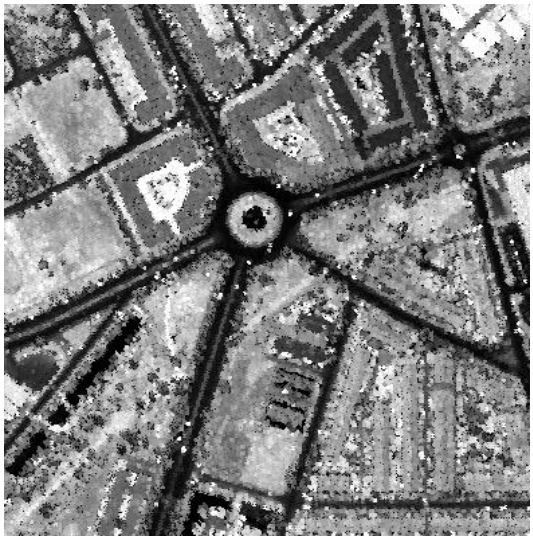
3.1. Image segmentation

Segmentation consists of group pixels that have similar properties; these pixels are treated as a set. The object of each segmentation is to simplify the appearance of an image in order to make it more meaningful and easy to analyse. Several methods can be used for segmenting: clustering, thresholding, or region growing. In this work, thresholding is the selected method used to carry out the segmentation. The region growing segmentation

method is applied to a single band. However, the images used in this work are composed of several bands. To consider a single band a Principal Component Analysis (PCA) was applied on four bands of the aerial image (RGBNIR). We considered only the first component of PCA.



(a)



(b)

Figure 3. (a) nDSM and (b) LIDAR intensity extract from the LIDAR data

3.2. Decision index

Four decision indexes have been used to detect the existing roads in the study area, as follows:

NDVI: the Normalized Difference Vegetation Index (NDVI) (1) is useful to determine the quantity, quality, and vegetation growth from aerial images taken by remote sensing. In the NDVI index, vigorous vegetation has high values. The NDVI index has widespread use in different remote sensing applications.

$$NDVI = \frac{NIR - RED}{NIR + RED} \quad (1)$$

LIDAR intensity: in addition to providing information related to high data, LIDAR data provides other information, such as the intensity of the return pulses. Intensity values are low in rough surfaces (with diffuse reflection), such as paved areas. This behavior is used to detect roads in the studied areas.

nDSM: Normalized Digital Surface Model (nDSM) is obtained from the digital surface model (DSM) and the digital terrain model (DTM). In the nDSM, every manmade construction in the field is represented. This model is very useful for detecting and differentiating land constructions and tall elements.

NDSPI: Normalized Difference Swimming Pool Index (NDSPI) (2) is an index created by the authors of this work. This index allows for the detection the swimming pools in an urban area. Roads and bodies of water have a similar response in the LINDAR variable of intensity. The NDSPI index allows for differentiating between roads and swimming pools.

$$NDSPI = \frac{BLUE - RED}{BLUE + RED} \quad (2)$$

3.3. Evidence theory

Once the four indexes are calculated, decision thresholds are set for each index and the evidence theory is applied.

The mathematical theory of evidence is a field in which the data sources are treated separately and their contributions are combined to provide a joint inference on the correct label for a pixel. This theory does not require a full probability model to work against the requirements of other approaches. It tries to benefit from the utilization of sets of assumptions or hypotheses rather than to address an hypothesis or an assumption separately, as is done in other approaches.

Evidence theory is applied to each of the segmented image regions and to every pixel that belongs to the original image. The result of both forms is compared with a ground truth prepared by the authors, in order to determine which of the methods offers the best results.

4. RESULTS

The results obtained at the pixel level and at the object level are showed below in Fig. 5 (a) and (b), respectively. The white color represents the roads and the black areas correspond to the background (Fig. 4).

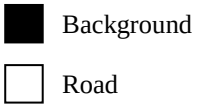
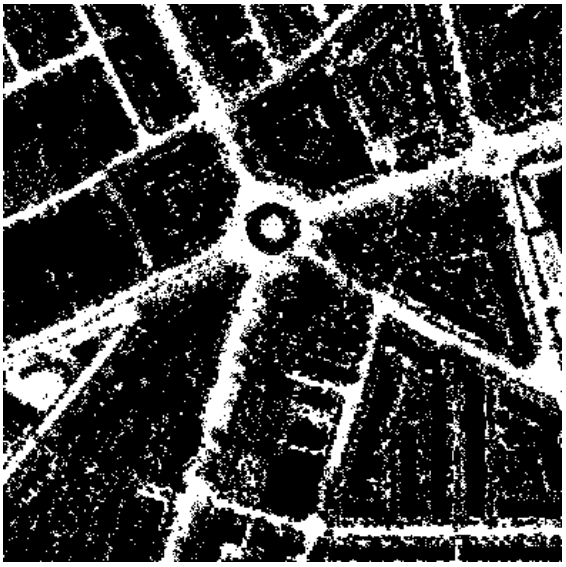


Figure 4. Legend of the categories



(a)

(b)

Figure 5. (a) Road detection at the pixel level and (b) at the object level.

One can observe that the pixel-level method provides a result that shows more noise than the object-level method. This is because the level pixel analyzes all the pixels in the image, one-by-one, while at the object level, the pixels sets are analyzed using common characteristics.

Apparently the object level results are better than the results obtained at the pixel level, but this is a subjective conclusion and it is necessary to confirm this numerically. To evaluate the results obtained at both levels, a ground truth has been produced by the authors of this work (Fig. 6). This was possible due to the fact that the studied region is well known by the authors because it is located close to the university where their research group works. Ground truth was used to analyze the results obtained from both methods. These results are shown in the confusion matrixes presented in Tab. 1 and 2.

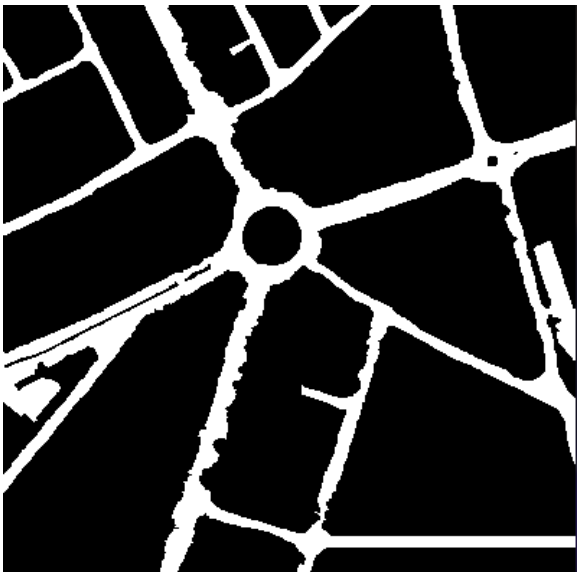


Figure 6. Ground truth used to evaluate the results obtained at the pixel level and the region level.

Table 1 – Confusion matrix of pixel level and results

	Roads	Background	Total pixels
Roads	22,250	15,588	37,838
Background	6,142	11,5901	122,043
Total	28,392	13,1489	159,881
Accuracy	**86.41**	Kappa index	**0.5884**
Omission	21.63	11.85	
Comission	41.20	5.03	

Table 2 – Confusion matrix of object level and results

	Roads	Background	Total pixels
Roads	22,724	13,729	36,453
Background	5668	117,760	123,428
Total	28,392	131,489	159,881
Accuracy	**87.87**	Kappa index	**0.6262**
Omission	19.96	10.44	
Comission	37.66	4.59	

As can be seen in the confusion matrixes, the accuracy obtained at the object level is higher than the results provided from the pixel level method. While the pixel level method obtained an accuracy of 86.41%, the object level method obtained an accuracy of 87.87%. The Kappa index is higher at the object level and all errors, both of omission and commission, are higher for the pixel-level object. Analyzing the quantitative results, it can be concluded that the object-level method provides better results than the pixel-level approach.

In order to compare the results obtained from the automatic classification method with other classification algorithms, a SVM supervised classification was carried out. When performing the classification, five classes are taken into consideration (just as in the automatic method), but four of those classes are grouped (bare soil, vegetation, building, and swimming pools) into a call background so that the final classification has two categories: roads and background. Fig. 7 shows the results of the SVM classification and Tab. 3 shows the results of the confusion matrix, which can be used to quantitatively analyze the classification that was conducted.

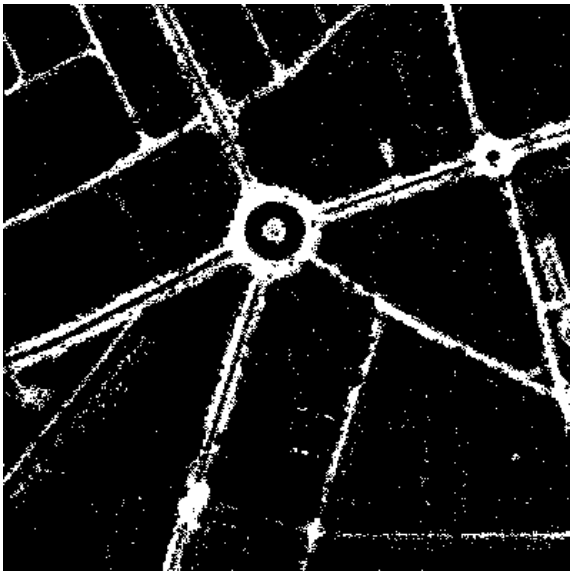


Figure 7. Road detection using SVM classifier.

Table 3 – Confusion matrix of SVM classification and results

	Roads	Background	Total pixels
Roads	14,574	2293	16,867
Background	13,818	12,9196	14,3014
Total	28,392	13,1489	15,9881
Accuracy	**89.92**	Kappa index	**0.5897**
Omission	48.67	1.74	
Comission	13.59	9.66	

As shown in Tab. 3, the accuracy obtained with the SVM supervised classification is 89.92%, which is more than 2% better than the accuracy obtained with the automatic region detection method. However, the Kappa index is higher for the object-level method results. Looking closer at the confusion matrix, it is found that the SVM classification provides 13,818pixels classified as background, when in fact that corresponds to the road pixels. This method properly classifies 14,574 road pixels. The automatic method correctly classifies 22,724 pixels as roads while only 5,668 pixels are labeled as background when in reality they represent a road. Moreover, the supervised classification properly labeled 129,196 background pixels, while the proposed method correctly classifies 131,489 pixels in this land cover. For this reason, the SVM classification has higher accuracy than the automatic method; however, it has been observed that the proposed detection method more correctly and accurately detects roads.

5. CONCLUSIONS

In this paper, we propose an automatic method for detecting roads in an urban environment from LIDAR data and aerial images. The method has been applied to the pixel level and the object level, and it was found that better results are obtained when the object level works. This is because those regions are sets of pixels that have common characteristics, which reduced the noise obtained by performing a pixel-level classification. An SVM supervised classification was also conducted and it has been observed that while the classification accuracy of that method is higher, the road detection is more accurate with the proposed method.

We conclude that in the proposed method, the objects are obtained at slightly better than pixel level, and that the quality of these results is comparable to a supervised SVM classification.

ACKNOLEDGMENT

The authors wish to thank the Spanish Ministerio de Ciencia e Innovación for financial support; project number CGL2010-15357.

6. REFERENCES

1. Wang, L., Wu, Y. & Yuan, J. (2010). Vegetation analysis of Hebei Province Using SPOT-VEGETATION NDVI", *International Conference on Geoinformatics.*

2. Huete, A.R. (1988). A Soil-adjusted vegetation index (SAVI), *Remote Sensing of environment*, 25, pp295-309.

3. Gómez Nieto, I. & Martín Isabel, M.P. Estudio comparativo de índices espectrales para la cartografía de áreas quemadas con imágenes Modis.

4. Swain, P.H. & Hauska, H. (1977). The decision tree classifier: design and potential, *IEEE Transaction on Geoscience Electronics*, vol. 15, pp142-147.

5. Rottensteiner, F. & Briese, Ch.(2003). Automatic generation of building models from lidar data and the integration of aerial images, *ISPRS*, vol. 34.

6. Chen, G., Hu, C. & Chen, H.(2011), Change detection of land use and land cover of the southeast coastal area in Fujian, *China*, *International Conference Multimedia Technology (ICMT)*, pp3456-3459.

7. Xu, H., Yang, M. & Liang, L. (2010). An improved random decision trees algorithm with application to land cover classification, *International Conference on Geoinformatics*.

8. Dempster, A.P., Laird, N.M. & Rubin, D.B. (1977). Maximum likelihood from incomplete data via the EM algorithm, J.R. Statist Soc., 39(1), pp1-38.

9. Hse, C.W. & Lin, C.J. (2002). A comparison of methods for multiclass support vector machines, IEEE Transaction on Neural Networks, 13, pp415-425.

10. Atkinson, P.M. & Tatnall, A.R.L. (1997). Neural networks in remote sensing, International Journal of Remote Sensing, 18(4), pp699-709.

11. He, L.M., Kong, F.S. & Shen, Z.Q. (2005). Multiclass SVM based land cover classification with multisource data, Proceedings of the Fourth International Conference on Machine Learning and Cybernetics, pp3541-3545.

12. Tan, K.C., Lim, H.S. & Mat Jafri, Z. (2011). Detection of Land Use/Land Cover Changes for Penang Island, Malaysia, Proceeding of the IEEE International Conference on Space Science and Communication, pp152-155.

13. O'Hara, C.G., King, J.S., Cartwright, J.H. & King, R.L.(2003). Multitemporal Land Use and Land Cover Classification of Urbanized Areas Within Sensitive Coastal Environments, IEEE Transactions on Geoscience and Remote Sensing, 41(9), pp2005-2014.

14. Holmes, W.S. (1966). Automatic photointerpretation and target location, Proceedings of the IEEE, 54, pp1679-1686.

15. Duda, T., Canty, M. & Klaus, D. (1999). Unsupervised land-use classification of multispectral satellite images. A comparison of conventional and fuzzy-logic based clustering algorithms, Proceedings of IEEE International Geoscience and Remote Sensing Symposium (IGARSS), pp. 1256-1258.

16. Bischof, H., Schneider, W. & Pinz, A.J.(1992). Multispectral classification of Landsat-images using neural networks, IEEE Transactions on Geoscience and Remote Sensing, 30, pp482-490.

17. Mena, J.B. & Malpica, J.A. (2005). An automatic method for road extraction in rural and semi-urban areas starting from high resolution satellite imagery, Pattern Recognition Letters, 26(9), pp1201-1220.

18. Mena, J.B. (2003). State of the art on automatic road extraction for GIS update: a novel classification, Pattern Recognition Letters, 24, pp3037-3058.

19. S. Myint, S., Gober, P., Brazel, A., Grossman-Clarke, S. & Weng, Q. (2011). Per-pixel vs. object-based classification of urban land cover extraction using high spatial resolution imagery, Remote Sensing of Environment, 115, pp1145-1161.

20. Raber, G.T., Jensen, J.R., Schill, S.R. & Schuckman, K. (2002). Creation of Digital Terrain Models Using an Adaptive Lidar Vegetation Point Removal Process, Photogrammetric Engineering & Remote Sensing, 68, pp1307-1315.