ROAD REGION DETECTION IN URBAN AREAS COMBINING HIGH-RESOLUTION RGB IMAGE AND LASER SCANNING DATA IN A CLASSIFICATION FRAMEWORK

A. P. Dal Poz[*], T. S. G. Mendes

Dept. of Cartography, São Paulo State University, R. Roberto Simonsen, 305, Presidente Prudente-SP, Brazil,
aluir@fct.unesp.br; tati_sussel@yahoo.com.br

Commission III

ABSTRACT:

This paper addresses the problem of road region detection in urban areas using an image classification approach. In order to minimize the spectral superposition of the road (asphalt) class with other classes, the Artificial Neural Networks (ANN) image classification method was used to classify geometrically-integrated high-resolution RGB aerial and laser-derived images. The RGB image was combined with different laser data layers and the ANN classification results showed that the radiometric and geometric laser data allows a better detection of road pixel.

1. INTRODUCTION

Various methods have been proposed for classification of remote sensing images taken from complex urban regions (Bellens et al., 2008, Pálsson et al., 2012). However, as emphasized by Bellens et al. (2008), the use of spectral data alone to classify this type of data frequently results in the overlapping of information classes. A possible strategy to minimize this problem consists in integrating airborne laser scanning (ALS) data and multispectral images, thereby improving the classification results.

Image classification methods have been used to detect roads in complex urban areas that, in some cases, serve as an initial step for more complex tasks regarding the road reconstruction. Few examples are presented in the following. Benkouider et al. (2011) developed a method to separate the road class in RGB SPOT images using spectral characteristics of roads in a classification process using Artificial Neural Networks (ANN), followed by a morphological post-processing to regularize the previously classified roads. A method for extracting roads from multispectral IKONOS images was proposed by Gao and Wu (2004). First, an unsupervised classification is applied to the IKONOS images, followed by post-processing steps applied to the resulting road class, including noise removal, skeletonization to extract road segments, and linking of road segments. Mancine et al. (2009) used the machine-learning algorithm AdaBoost (i.e., Adaptive Boosting) to classify multispectral aerial images that were integrated with ALS data. The Hough transform was then used to reconstruct roundabouts and roads.

In this paper we propose a study of different combinations between optical images and geometric and radiometric ALS data using the ANN classification method. Although several information classes (grass, trees, building roofs, asphalted roads etc.) are defined in the classification problem, our main focus is on the detection of the road pixels. The present article is organized as follows. The proposed method is introduced in Section 2, the experimental results and discussions are presented in Section 3, and the major conclusions are presented in Section 4.

2. METHOD

Our proposed study on the feasibility of using ALS data combined with optical images in a classification framework, using ANN, addresses two main steps: 1) geometric integration of ALS data and optical images and pre-processing; and 2) ANN classification involving different combinations of geometrically-integrated optical and ALS-derived images, with main focus on the separation capability of the road class.

2.1 Data integration and pre-processing

The ALS data and aerial images can be integrated by projecting the ALS point cloud onto the aerial image followed by a regularization of the projected points according to a grid with the same resolution of the aerial image. The ALS point cloud projection onto the image space involves several mathematical procedures, as e.g. one based on the collinearity equations. Details on these mathematical procedures can be found in relevant literature, as e.g. in Wolf and Dewitt (2000). The ALS points in image space have row and column coordinates of the image coordinate system (ICS), plus respective orthometric heights (h) and laser pulse intensities (I).

The photogrammetric procedure mentioned above allows for the generation of both the Digital Surface Model (DSM) and the Digital Terrain Model (DTM) in the ICS. A third grid, the normalized DSM (DSMn), is generated based on a subtraction of the DTM from the DSM. The ground and aboveground objects can be separated in the DSMn by a simple thresholding process taking into consideration that aboveground objects (like vegetation and building) have a minimum height. It is expected that the DSMn with thresholding can supply a hard constraint to avoid classifying aboveground objects as roads.

As each ALS point has a radiometric response (i. e., the so called laser pulse return intensity), together with the X, Y, and h

---

[*] Corresponding author.

coordinates, the same procedure described above for the DSM and DTM generation in the ICS can be used to generate a laser pulse intensity image (or shortly referred to as the intensity image), also in the ICS and with same resolution of the input aerial image. Roads in intensity images are easily identified because they possess homogeneous spectral responses and appear very dark due to the asphalt's low reflectivity of the laser pulse (approximately 17% according to Wehr and Lohr (1999)). Another significant characteristic of the intensity image is that it is not affected by shadows of aboveground objects as in optical images. This means that a road segment obstructed by shadow in an aerial image can be seen in an intensity image. Therefore, the intensity image can contribute considerably to refining information during the classification of images. Despite these advantages, the intensity images exhibit certain weaknesses. The first of these is the noise characteristic that directly affects the results of the classification. This issue can be minimized by applying a median filter or a morphological filter. An additional problem with this type of data is the presence of objects that show radiometric responses similar to those of roads, such as the building roofs and certain types of vegetation.

The procedure discussed above generates the following geometrically-integrated layers: R, G, and B image layers; DSMn segmented using the thresholding method (the threshold used was 2.5 m), and the intensity image smoothed using the median method. The DSM and DTM used to derive the DSMn were generated as follows. The nearest neighbor interpolation method was used for deriving the DSM. Points on the ground were selected visually in the laser intensity image and used to generate the DTM by using the kriging interpolation method. As already mentioned, the DSM and DTM were generated in the ICS at the same resolution of the RGB image.

2.2 Supervised classification through ANN

The data layers discussed in Subsection 2.1 are used to isolate the road class by employing the ANN image classification method (Gonzalez and Woods, 2008). In order to model the scene as a whole, the following additional classes are defined: grass, trees, buildings with gray roofs, buildings with red roofs, and concrete.

The backpropagation algorithm (Richards and Jia, 2006) was used to train the networks. The input layers were combined in different ways to determine the best combinations for separating the road class and to verify the effective contribution of input layers from the ALS data. All of the input data layers in the network were normalized in the range 0 to 255. Table 1 shows examples of relevant combinations of data layers for evaluating the contribution of ALS data combined with a RGB image into the classification procedure by ANN.

Combination	Input Layer Data
RGB	R,G, and B
RGB+DSMnThr	R, G, B, and DSMn with thresholding (DSMnThr)
RGB+IS	R,G, B, and intensity image with median smoothing (IS)
RGB+DSMnThr+IS	R,G, B, DSMnThr, and IS

Table 1. Examples of combinations of input data layers for classification by ANN

The first combination (RGB) only utilizes layers of the RGB image and is used to analyse and compare the contributions of

the subsequent additions of the ALS data during the classification process. The second combination (RGB+DSMnThr) adds the geometric ALS data to the RGB image layers by using the image representing the DSMnThr in the ICS. The major objective of performing this combination was to check the hypothesis by which the geometric ALS data is useful for separating classes corresponding to aboveground objects (for example, roofs and trees) from those that are on the terrain surface. The intensity image in the ICS was smoothed using the median filter (resulting in the IS image) and added to the RGB image layers, giving rise to the third combination (RGB+IS) in Table 1. The rationality of smoothing the intensity image is that it is very noisy. The IS image is expected to be useful in the classification process because roads are very well contrast in it. Moreover, due to the capability of the laser pulse in penetrating the tree structures and reaching the ground, road obstructed by trees can be totally or partially visible in the IS image. In order to verify the joint contribution of the geometric and radiometric ALS data, when combined to the RGB image, the last combination (RGB+DSMnThr+IS) mixes the RGB image layers, the DSMnThr image layer, and IS image layer.

3. EXPERIMENTAL RESULTS

The test area comprised an urban region of the city of Curitiba, Southern Brazil, for which an aerial high-resolution (GSD ~ 0.2 m) RGB image and an ALS point cloud at an average resolution of 0.5 m were available.

Classifications using ANN, which were performed with the IDRISI Andes software (Clark Labs, Worcester, MA, USA), tested different network architectures. A learning rate of 0.01 and a momentum factor of 0.5 were used to optimize the convergence of the network. In order to verify the influence of the ALS data layers in the classification process by ANN, several input layers were combined, as shown in Table 1. Results are presented for the aerial sub-image shown in Figure 1.

Figure 1. RGB aerial sub-image

In order to obtain the classification results for the RGB image (Figure 2), the ANN was trained for about 10% of the data based on an architecture with a input layer containing three neurons (R, G, and B image layers), two hidden layers containing twenty one neurons and forty neurons, and a output layer containing six neurons (information classes). The used architecture (in short, 3-21-40-6) was selected by the try and error method based on the best OA (i. e., 71%). The worst

results were achieved for the concrete and gray roof classes (corresponding producer accuracies were below 40%). On the other hand, the best result was found for the red roof class, for which the producer accuracy was 90%. Although relatively satisfactory producer accuracy was achieved (85% producer accuracy) for the road class, it is possible to note that the road class was confused with concrete and gray roof classes. Also, large building shadows were classified as roads. Due to the above reason, the kappa coefficient was very low, i. e., 65%.
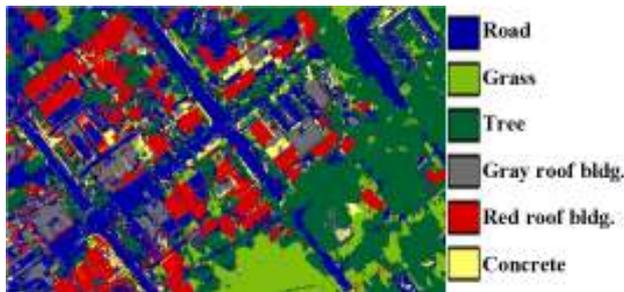

Figure 2. Results obtained for the RGB image

Figure 3 shows the classification results for the RGB+DSMnThr combination. The ANN was trained similarly to the previous example, resulting in a 4-26-46-6 architecture with 86% OA, which is considerably better than the result obtained using only the RGB image. Although the results obtained with the concrete and gray roof classes continued to be the worst, corresponding OAs ranges now from 60% to 70%. Particularly, confusions between road class and other spectrally similar classes in the RGB image (i. e., gray roof and concrete) were reduced. Please note that, for instance, gray roofs were not classified as road and the misclassification of shadows as roads were considerably reduced. It is also worthy to note that roads obstructed by trees, nearly the bottom right-hand corner of the image, were partially detected. As a result, the road producer accuracy (96%) and the kappa coefficient (79%) were much better than ones obtained for the RGB classification.
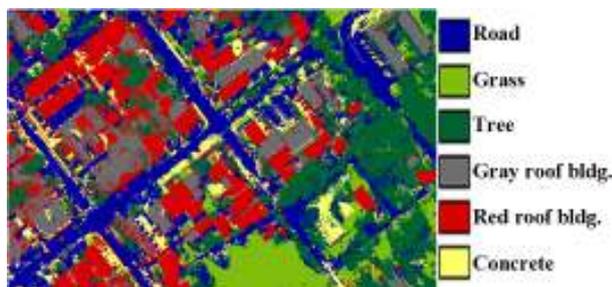

Figure 3. Results obtained for the RGB+DSMnThr combination.

Next example (Figure 4) combines the RGB image with the ALS intensity image with a median smoothing (IS). The adopted architecture (88% OA) for this combination was 4-45-15-6, which was trained as in previous examples. The classification results are similar to ones obtained by using the DSMnThr image in almost all aspects, meaning that the intensity ALS data is also useful for avoiding the confusion

between the road class and other spectrally similar classes in the RGB image (like gray roofs). The main observable difference in relation to the previous combination (RGB+DSMnThr) is that parking lots were better detected (as road pixels) using the IS image. As a result, the road producer accuracy was slightly worse (94%) for this test. The kappa coefficient was about 75%.
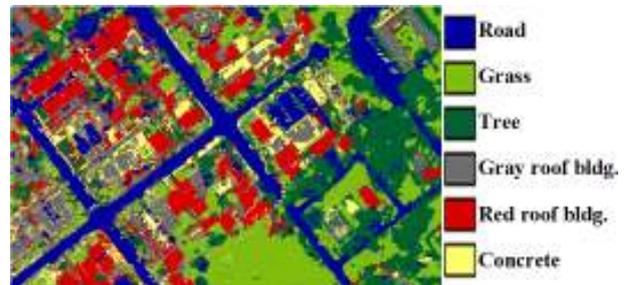

Figure 4. Results obtained for the RGB+IS combination

Next combination (RGB+DSMnThr+IS) mixes the three layers from the RGB image, one layer from the geometric ALS data (DSMnThr image), and one layer from the radiometric ALS data (IS image). The results obtained (Figure 5) for this combination were based on the architecture 5-30-50-6, which provided the best OA (90%). The kappa coefficient was 80%, which was the better one. In general, the classification results were better for all classes (For instance, the producer accuracies were 75%, 92%, and 87% for the gray roof, red roof, and grass classes, respectively), but it is worthy to emphasize the very high producer accuracy (96%) obtained for our interest class, i. e., the road class. In fact, Figure 5 shows that roads were successfully classified in regions obstructed by trees and few pixels were wrongly classified as road inside the blocks, excluding those related to parking lots having asphalt material.
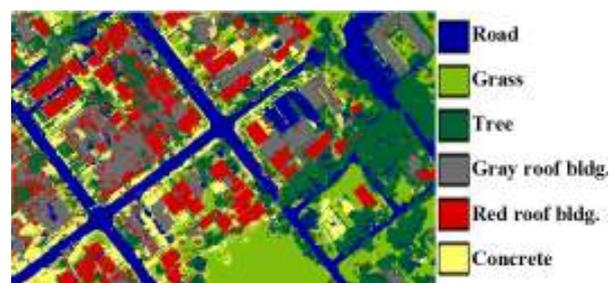

Figure 5. Results obtained for the RGB+DSMnThr+IS combination

4. CONCLUSIONS AND FUTURE WORK

This paper described an experimental study with the main focus on the detection of road pixels based on an ANN classification procedure of geometrically-integrated high-resolution RGB aerial and laser-derived images.
The classification results showed that the RGB image provided the worst performance due to well-known confusions among objects showing similar spectral responses, as e. g. those related to roads and gray roofs. Mainly the combination among the

RGB image and radiometric and geometric ALS data clearly demonstrated that ALS data improved the classification as a whole and, particularly, the classification of our class of interest (i.e., roads).

Our future work will be focused on the development of post-processing techniques to refine the detected road regions and to reconstruct the road network using the refined road regions.

ACKNOWLEDGEMENTS

REFERENCES

Bellens, R.; Gautama, S.; Martinez-Fonte, L.; Philips, W.; Chan, J. C.; Canters, F., 2008. Improved Classification of VHR Images of Urban Areas using Directional Morphological Profiles. *IEEE Transaction on Geoscience and Remote Sensing*, 46(10), pp. 2803-2813.

Benkouider, F.; Hamami, L.; Abdellaoui A., 2011. Use of the Neural Net for Road Extraction from Satellite Images, Application in the City of Laghouat (Algeria). In: *Progress in Electromagnetics Research Symposium 2011,* Vol. 7, pp. 146-150.

Gao, J.; Wu, L., 2004. Automatic Extraction of Road Networks in Urban Areas from Ikonos Imagery Based on Spatial Reasoning. In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Istanbul, 2004, Vol. XXXV, pp. 331-337.

Gonzalez, R. C., Woods, R. E., *Digital Image Processing.* 3th Edition. Prentice Hall, 2008, 954 p..

Mancine, A.; Frontoni, E.; Zingaretti, P. 2009. Automatic extraction of urban objects from multi-source aerial data. In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Paris, Vol. XXXVIII, pp. 13-18.

Pálsson, F.; Sveinsson, J. R.; Benediktsson, J. A.; Aanæs, H., 2012. Classification of Pansharpened Urban Satellite Images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 5(1), pp. 281-297.

Richards, J. A.; Jia, X., *Remote Sensing Digital Image Analysis.* 4th Edition Berlin-Heidelberg. Springer-Verlag, 2006, 439 p..

Wehr, A.; Lohr, U., Airborne LASER scanning – an introduction and overview, *ISPRS J. Photogramm. Remote Sens.*, v. 54, n.. 2-3, pp. 68 - 82, 1999.

Wolf, P. R.; Dewitt, B. A., 2000. *Elements of Photogrammetry: with applications in GIS*. Boston: McGraw-Hill, 608 p.