Remote Identification of Housing Buildings with High-Resolution Remote Sensing*

José Luis Silván-Cárdenas[1], Juan Andrés Almazán-González[1], and Stéphane A. Couturier[2]

[1] Centro de Investigación en Geografía y Geomática "Ing. Jorge L. Tamayo" A.C. Contoy 137, Lomas de Padierna, Tlalpan, Mexico D.F. 14240
jlsilvan@centrogeo.org.mx
http://www.centrogeo.org.mx

[2] Laboratorio de Análisis Geoespacial, Instituto de Geografía, UNAM Cto. de la Investigación Científica s/n, CU, Coyoacán, México D.F. 04510
andres@igg.unam.mx
http://www.igeograf.unam.mx/sigg/

Abstract. Identifying housing buildings from afar is required for many urban planning and management tasks, including population estimations, risk assessment, transportation route design, market area delineation and many decision making processes. High-resolution remote sensing provides a cost-effective method for characterizing buildings and, ultimately, determining its most likely use. In this study we combined high-resolution multispectral images and LiDAR point clouds to compute building characteristics at the parcel level. Tax parcels were then classified in one of four classes (three residential classes and one non-residential class) using three classification methods: Maximum likelihood classification (MLC), Suport Vector Machines (SVM) with linear kernel and SVM with non-linear kernel. The accuracy assessment from a random sample showed that the maximum MLC was the most accurate method followed by SVM with linear kernel. The best classification method was then applied to the whole study area and the residential class was used to mask-out non-residential buildings from a building footprint layer.

Keywords: Remote sensing, LiDAR, housing units, land use classification.

1 Introduction

Mapping housing units in urban environments is needed for a number of urban planning process as well as for scientific studies such as population estimation, risk analysis and evacuation routes design. Identifying housing units from remote sensing involves building detection and land use classification with at least two classes, one for residential and one for non-residential land uses.

* This research was funded by the National Council of Science and Technology (CONACYT) project "Small-area population estimation by means of remote sensing" Grant 187593.

J.F. Martínez-Trinidad et al. (Eds.): MCPR 2014, LNCS 8495, pp. 380–390, 2014.

The land use classification problem has been addressed through observable land cover characteristics from remote sensing products, such as spectral reflectance, texture, impervious surface and surface temperature, as well as from existing geographic information layers, including zone maps, population density and road maps, among others [1]. A wealth number of decision rules have been also developed and applied to land use/land cover classification problems, among which the maximum likelihood decision rule stands as the most common one. In the last two decades, statistical learning theory has provided non-parametric decision rules for solving complex classification problems. These techniques include artificial neural networks (ANN), decision trees (DT) and support vector machines (SVM). In general, non-parametric methods tend to outperform the maximum likelihood classification (MLC) method with varying degree of performance depending on specific configurations, with SVM being less sensitive to the increase in data dimensionality and presence of correlation [2–4].

Despite this, observations are not conclusive because most comparative studies have applied decision rules on pixels (pixel-based classification) rather than on segments extracted from the image (object-based classification), or on preestablished administrative boundaries (field-based classification). While pixel-based classification methods tend to be more efficient as they can operate exclusively with data in raster format, some researchers have also noted that classification results are noisy and less accurate than when classifying groups of pixels altogether, specially when using high-resolution data [5]. Furthermore, in urban environments, tax parcels are owner's land demarcations and, therefore, they has been claimed as the natural land use unit [6, 7].

This article describes mayor processing steps of a methodology for identifying housing units by integrating LiDAR altimetry data, multispectral satellite images and cadastral layers from a GIS. Since the building detection problem has been the focus of a related study [8] this article focus on evaluation of the land use classifications component of the whole process. The rest of the paper is organized as follows. Section 2 presents the fundamentals of Support Vector Machines (SVM), which is one of the methods tested in the study. Then, in Section 3 the overall processing flow for extracting housing units is described. In Section 4 the accuracy assessment results for the land use classification are presented and its impact in the whole process are discussed. Major conclusions drawn from the study are presented in Section 5.

2 Background

The theory of support vector machines (SVMs) formulates linear discriminants for solving binary classification problems where classes may overlap in the feature space [9]. The linear version of the problem consists in finding the maximum margin hyperplane that separates samples from two given classes. Mathematically, this is formulated as follows.

Given a training set of n exemplar pairs $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ of feature vectors $x_i \in \mathbb{R}^d$ and class labels $y_i \in \{1, -1\}$, find the hyperplane

$x_i^T\omega + \theta = 0$ that creates the biggest margin between the training points for the two classes. This problem is equivalent to the following optimization problem:

$$\min \frac{1}{2}||\omega||^2 + c\sum \varepsilon_i \quad \text{subject to} \quad y_i(x_i^T\omega + \theta) \geq 1 - \varepsilon_i, \varepsilon_i \geq 0, \forall i \tag{1}$$

where ε_i are slack variables that account for the overlapping of classes and c expresses the cost of accepting discrimination errors. This cost is infinity in the separable case.

The optimization problem above is quadratic with linear constraints. This convex problem is typically solved with Lagrange multipliers, which also results convenient because the solution depends on scalar products of certain feature vectors, called *support vectors*. Moreover, this allows for construction of nonlinear boundaries through the use of kernels that represent the scalar product in a transformed vector space. Most popular kernel functions include a polynomial of a given degree, the sigmoid function and the radial basis function (RBF). The latter, which was used in this study, is expressed as

$$K(x, x') = e^{-\gamma||x-x'||^2} \tag{2}$$

The extensions of SVMs to solve multi-class problems include methods that consider all classes at once and those that use multiple binary SVMs, whose outputs are combined through common strategies such as "one-against-all", "one-against-one," and directed acyclic graphs (DAG). Since solving multiclass SVM in one step implies a much larger optimization problem, experiments have shown that the "one-against-one" and DAG methods are more suitable for practical use than the other methods [10].

3 Materials and Methods

This section describes the major processing steps and datasets used in the study. The overall workflow is outlined in Fig. 1 and a 3-d view of some of the data sets generated are shown in Fig. 3.

3.1 Data Used

The study area is located in the historical and cultural center of Mexico City, and corresponds to one of the borough of the Federal district, Cuatémoc. Being the oldest part of the city, it hosts historical buildings as well as modern skyscrapers, government buildings, market and commercial centers, and residential buildings, most of which are multifamily type. With a total area of 32.44 km^2 and a population of over half million people as of 2010, it represents the highest population density of the city (16,000 peoples/km^2).

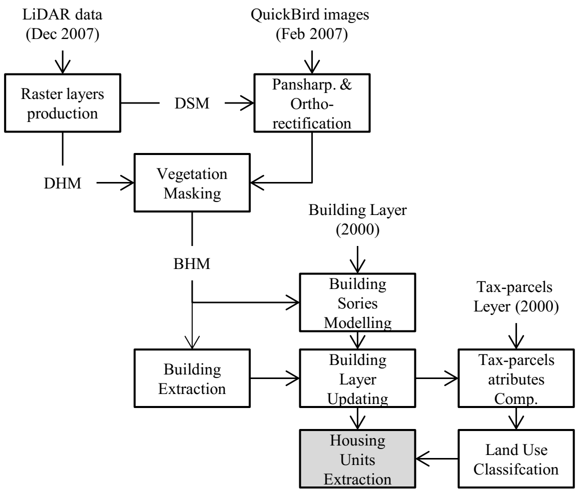

Fig. 1. Work flow for housing units identification

LiDAR Data and Derived Products. In November and December of 2007, the ALS50-II LiDAR system was flown over the Mexico valley covering the entire Federal District and the area of Chalco in the State of Mexico. The system delivered 3-d coordinates of locations where the laser hit the Earth's surface. Along with the 3-d coordinates, the system recorded the return intensity and the return number (for up to four returns per pulse) attached to each point. The average point density was 0.433 points/m^2 with average point distance of 3.0 m and a vertical accuracy of 7.3 cm (RMSz).

The data was provided in LAS format by the National Institute of Geography, Statistics and Informatics (INEGI) and then rasterized in tiles of 1 km by 1 km to form digital surface models (DSM) of 1 meter resolution (blocks of 1000-by-1000 pixels). A mosaic DSM of 7-by-8 km was also built to cover the entire study area. Each DSM tile was also normalised by removing the terrain variation using the ground filtering method described in [11], thus resulting in a digital height model (DHM) of the same resolution and coverage as the DSM (7-by-8 km).

Multispectral Imagery. A QuickBird bundle image with multispectral (at 2.4 m spatial resolution) and panchromatic (at 0.6 m spatial resolution) bands acquired on February 2007 was used in this study. The four multispectral bands were fused with the panchromatic band using the Gram-Schmidt pan-sharpening method [12] resulting in four multispectral bands of 1 m resolution. The spatially enhanced image was then orthorectified using the rational polynomial coefficients provided by the image vendor. This method required the DSM and a set of on-screen ground control points.

Cadastral Layers. Cadastral data with a time reference of year 2000 was acquired through the Ministry of Urban Development and Housing (SEDUVI). The data was acquired in DXF format, so that layers were extracted and coverted to ESRI's shapefile format. Two shapefiles were built for the entire administrative limits of Cuauhtémoc, one for tax parcels boundaries and one for building boundaries. Both layers underwent an extensive editing process to correct numerous errors, some of which were generated during the conversion process.

In addition to the geometry information, two annotation layers available in the original DXF files were also extracted, namely the number of stories and the land use key. The number of stories layer was included as field of the building geodatabase, whereas the land use key was included as field of the tax parcel geodatabase.

3.2 Building Stories Modelling

The number of stories of a building is strongly related to the buildings height. Building height was computed for each building footprint using a LiDAR-derived product. Although the DHM represents the building height for every square meter, it also contained some other information that can lead to erroneous estimation of building height, particularly the high vegetation hanging over the building roofs.

In order to eliminate vegetation height from the DHM, a vegetation mask was generated with the QuickBird image. This was accomplished by simply thresholding the normalized difference vegetation index (NDVI), so that a new raster surface, termed the building height model (BHM), was generated by setting height values in the DHM to zero for detected vegetation pixels. The building height was then calculated as the median value of BHM pixels within the building polygon. The median height was preferred over the mean as it is more robust to outliers due to above-roof installations such as antennas and water tanks.

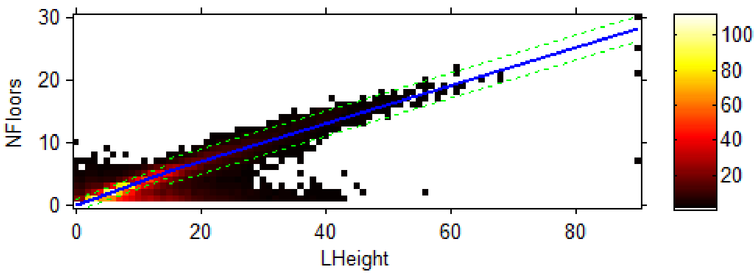

Fig. 2. Piece-wise linear model (blue line) fitted to the NFloors-LHeight scatterplot (ramp color image). The discontinuous green lines show a tolerance error of ±2 floors for buildings higher than 15 meters and of ±1 for lower buildings.

The scatter-plot of building height (LHeight in Fig. 2) against number of building stories (NFloors) was examined and a piece-wise linear model was visually adjusted. The model is expressed as:

$$N = \begin{cases} [H/2.8] & \text{for } H < 15 \\ [H/3.3] & \text{for } H \geq 15 \end{cases} \tag{3}$$

where N and H represent the number of building stories and the corresponding building height, respectively.

3.3 Building Extraction

A method for extracting building footprints was also implemented based on a previously developed method for tree-crown segmentation [13]. Two major differences were implemented in the new method.

First, instead of taking the original DHM as input, the gradient of the BHM was used as input. The rationale of this method is that the variations in building roofs height naturally creates edges between building roofs. This is not necessarily true when a number of attached buildings have the same height. Nonetheless, the tax parcel polygons can be used to further split those segments that covered more than one tax parcel.

Second, once potential building segments were extracted, segments with a median height lower than a minimum height (2.0 m) were eliminated from the segmentation image. Also, segments with an area lower than a minimum area (e.g., 16m^2) were merged to the largest neighbouring segment.

The segments so computed are used to select those buildings that required updating. So far the updating has been carried out manually, while automated methods are being developed. At this point, the height of all buildings was also updated by applying the piece-wise linear model of Eq. 3. In order to account for errors in building height estimations, the updating took place only if the estimated number of stories differed from the original value beyond one 1 floor for buildings lowers than 15 meters and of 2 floors for higher buildings. This criterion is based on the observation that taller buildings are less unlikely to change by a few levels.

3.4 Land Use Classification

A number of tax-parcel attributes fields were derived at the parcel level by aggregating building features at parcel level. These fields included the maximum number of building stories, the total built-up area, the number of buildings in the parcel, percent of built-up area, the total area of the parcel, and the habitable space. The latter is a volumetric measure of the potentially habitable space of buildings within a parcel and is defined as:

$$S_i = \sum_j A_{i,j} N_{i,j} \tag{4}$$

where $A_{i,j}$ and $N_{i,j}$ are the area and number of stories of the i-th building in parcel j.

Three supervised classification methods were tested. The first classification method is the well known maximum likelihood classification (MLC) method, which is a parametric method that depends on second order statistics of Gaussian probability distribution functions (PDF) [14]. It is the optimal solution to a classification problem when the PDF of classes is multivariate Gaussian. The second and third methods corresponded to SVM with linear and with nonlinear kernels, respectively.

Both the linear and nonlinear formulations were implemented in MATLAB based on previously published codes [15]. However, since SVM has been designed for solving binary classification problems, a majority voting of the One-Against-One approach was also implemented as suggested by [10]. According this approach, one SVM is developed for each pair of classes, whereas the classification result from each SVM is treated as a 'vote' to one of the classes. Then, the class assigned to each input feature vector is based on the majority of votes. When ties result, the SVM involving the tied classes has a quality vote that is used to solve the tie. If tie cannot be solved even with the quality vote, then the feature is tagged as unclassified.

4 Results

Despite the detailed classification system that was available through the land use classification key, a simplified classification system was used in this study as provided in Table 1. This classification system obeyed the objective of examining the potential of supervised classification methods to discriminate residential land uses from a reduced number of parcel and building characteristics, namely the total area of the parcel, the maximum number of stories, the built-up area, the number of buildings in the parcel, percent of built-up area, and the habitable space.

Table 1. Land use classification system used in the study

Level 1	Level 2	Description
	R0	No buildings
Residential	R1	Buildings with 1 or 2 stories
	R2	Buildings with 3 or more stories
Non Residential	NR	Commercial, Industrial and all other

A random sample was selected for training and testing purpose from the entire set of tax parcels. In order to ensure the representativeness of the various residential and non-residential land uses a two-stages strategy was implemented. In the first stage a mesh of cells of 9 hectares (300 by 300 m) was overlaid on the study area. The cells were ranked according to a diversity index, which accounted for the number of distinct land uses present within the cell. The first 20 cells with the highest diversity index were used to narrow down the potential tax parcels that were going to be used. In the second selection stage, a stratified random selection was applied, according to which 100 tax parcels were selected for the classes R1, R2 and NR described in Table 1. Whilst there were only 30 tax parcels of the class R0 in the entire study area, all of which were included in the sample. Finally, the sample was randomly split in two sets with half-half proportion except for R0, in which all 30 samples were included in both sets.

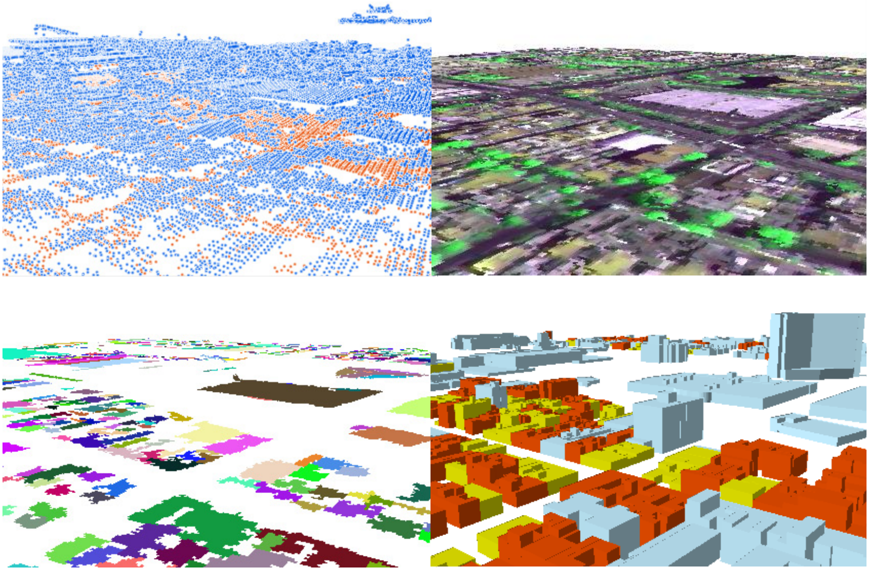

Fig. 3. 3-d view of major products derived from the study: classified LiDAR point cloud (upper left), false calor composite of QuickBird image wrapped on the DTM (upper right), Segments extracted from the BHM (lower left) and extruded polyhedral representation of MLC-classified buildings (lower right). Residential buildings are in yellow (R1) and red (R2), whereas non-residential are shown in light blue.

The three classification methods were trained using one of these sets (training set), whereas classification accuracy was assessed using the other set (testing set). The confusion matrices are shown in tables Table 2. Results indicated that MLC was the most accurate classification method followed by the linear SVM and non-linear SVM with an overall accuracy of 73.3%, 63.3% and 60.5%, respectively. The linear SVM presented a couple of tie conflicts that could not be resolved with the quality vote criterion. Nonetheless, even if these conflicts were solved in one or another sense, the accuracy of the classifier remained under that of MLC. When looking the per-class user and producer accuracies, MLC showed consistently the best performance for all classes whereas all methods showed the best performance for the class R0 (User Accuracy ranging from 70% to 90%)

and the worst performance for the class NR (User Accuracy ranging from 40% to 57%). This result suggests that while class R0 can accurately be represented by a Gaussian PDF, the other classes were not properly represented by this PDF. Furthermore, the relative low performance of SVM methods may have been affected by the voting strategy, the relatively small sample size, the overfilling of training sample.

Table 2. Confusion matrices for each classification method tested (a) MLC, (b) Linear SVM (c) Non-linear SVM

(a)

	Reference				
Class	R0	R1	R2	NR	Total
R0	30	0	0	3	33
R1	0	44	1	9	54
R2	0	2	37	17	56
NR	0	4	12	21	37
Total	30	50	50	50	180

(b)

	Reference				
Class	R0	R1	R2	NR	Total
R0	30	0	0	3	38
R1	0	39	1	9	52
R2	0	0	37	17	40
NR	0	7	12	21	48
Unclassified	0	2	0	0	2
Total	30	50	50	50	180

(c)

	Reference				
Class	R0	R1	R2	NR	Total
R0	30	2	0	11	43
R1	0	37	8	9	54
R2	0	6	25	13	44
NR	0	5	17	17	39
Total	30	50	50	50	180

The classification accuracy reflects the accuracy of correctly identifying housing buildings assuming buildings have been accurately extracted. This assumption is nearly true in our case, as an existing building layer was manually updated. If a fully automated updating method is to be implemented, then the building extraction method can further lower the overall identification accuracy. In a previous study [8] we have evaluated some building detection methods which can extract buildings with average detection rate of around 85% and with average commission errors of around 20%. With this figures, one may identify residential buildings with an accuracy as high as 60% which is only a moderate accuracy.

5 Conclusions

A method for identifying residential buildings based on remote sensing and GIS layers was outlined. The two major components of the method include building extraction, characterization and updating through LiDAR and multispectral imagery. The second part involved per-field land use classification of buildings. The accuracy assessment based on a stratified random sample showed that the MLC showed the highest accuracy over SVMs. This seems to contradict prior studies where SVM outperformed MLC [2–4]. However, it should be noted that those studies have performed the classification and its accuracy assessment at the pixel-level, while this study was based on tax parcels. Moreover, that the number tax parcel features have been kept relatively low (6 features) maybe the reason why MLC still performed well. Further investigation is needed to determine if SVM with a higher number of attributes, such as vegetation attributes, can lead to better performance than that achieved by low-dimensional MLC.

References

1. Lu, D., Weng, Q.: Use of impervious surface in urban land-use classification. Remote Sensing of Environment 102(1), 146–160 (2006)
2. Dixon, B., Candade, N.: Multispectral landuse classification using neural networks and support vector machines: one or the other, or both? International Journal of Remote Sensing 29(4), 1185–1206 (2008)
3. Guan, H., Li, J., Chapman, M.A., Zhong, L., Ren, Q.: Support vector machine for urban land-use classification using lidar point clouds and aerial imagery. In: International Symposium on LiDAR and RADAR Mapping: Technologies and Applications, Nanjing, China, pp. 26–29 (May 2011)
4. Srivastava, P.K., Han, D., Rico-Ramirez, M.A., Bray, M., Islam, T.: Selection of classification techniques for land use/land cover change investigation. Advances in Space Research 50(9), 1250–1265 (2012)
5. Myint, S.W., Gober, P., Brazel, A., Grossman-Clarke, S., Weng, Q.: Per-pixel vs. object-based classification of urban land cover extraction using high spatial resolution imagery. Remote Sensing of Environment 115(5), 1145–1161 (2011)
6. Bauer, T., Steinnocher, K.: Per-parcel land use classification in urban areas applying a rule-based technique. GeoBIT/GIS 6, 24–27 (2001)
7. Wu, S., Silván-Cárdenas, J., Wang, L.: Per-field urban land use classification based on tax parcel boundaries. International Journal of Remote Sensing 28(12), 2777–2801 (2007)
8. Silván-Cárdenas, J.L., Wang, L.: Extraction of buildings footprint from liDAR altimetry data with the hermite transform. In: Martínez-Trinidad, J.F., Carrasco-Ochoa, J.A., Ben-Youssef Brants, C., Hancock, E.R. (eds.) MCPR 2011. LNCS, vol. 6718, pp. 314–321. Springer, Heidelberg (2011)
9. Hastie, T., Tibshirani, R., Friedman, J.: 12 Support Vector Machines and Flexible Discriminants. In: The Elements of Statistical Learning, 2nd edn. Springer Series in Statistics, pp. 417–458. Springer (2009)
10. Hsu, C.W., Lin, C.J.: A comparison of methods for multiclass support vector machines. IEEE Transactions on Neural Networks 13(2), 415–425 (2002)
11. Silván-Cárdenas, J.L.: A multiscale erosion operator for discriminating ground points in lidar point clouds. In: Carrasco-Ochoa, J.A., Martínez-Trinidad, J.F., Rodríguez, J.S., di Baja, G.S. (eds.) MCPR 2012. LNCS, vol. 7914, pp. 213–223. Springer, Heidelberg (2013)
12. Laben, C.A., Brower, B.V.: Process for enhancing the spatial resolution of multispectral imagery using pan-sharpening (2000)

13. Silván-Cárdenas, J.L.: A segmentation method for tree crown detection and modelling from liDAR measurements. In: Carrasco-Ochoa, J.A., Martínez-Trinidad, J.F., Olvera López, J.A., Boyer, K.L. (eds.) MCPR 2012. LNCS, vol. 7329, pp. 65–74. Springer, Heidelberg (2012)
14. Richards, J.A.: 8-Supervised Classification Techniques. In: Remote Sensing Digital Image Analysis: An Introduction, 5th edn., pp. 247–318. Springer (2003)
15. Mangasarian, O., Musicant, D.: Lagrangian support vector machines. Journal of Machine Learning Research 1, 161–177 (2001)