Producing a Building Change Map for Urban Management Purposes

Teresa SANTOS[a1], Sérgio FREIRE[a], Ana FONSECA[b], and José António TENEDÓRIO[a]
[a] *e-GEO, Faculdade de Ciências Sociais e Humanas, FCSH, Universidade Nova de Lisboa, Av. Berna, 26–C, 1069–061 Lisboa;*
[b] *National Laboratory for Civil Engineering (LNEC), Lisbon, Portugal*

Abstract. The high rate of changes in cities requires the existence of matching geographic information in order to allow proper land monitoring and planning. The work presented in this paper concerns the exploration of very-high resolution (VHR) imagery as an alternative source of geospatial information for large scale mapping to assist municipal urban planning in Portugal. The goal is to produce an alarm system that indicates the location of potential changes in the built-up zones This layer can be used by the municipal technical staff as the basis for manual editing, following the technical specifications indicated for the 1:1 000 scale.

Keywords. QuickBird, LiDAR, feature extraction, change detection, Lisbon

Introduction

The municipal services require updated information mainly on three land cover classes: “Buildings”, “Roads” and “Green areas”. From these classes, buildings are the most relevant elements for the municipal activity, because there are administrative acts that occur on a daily basis, that require updated information. Situations like land subdivision into lots, redeveloped and/or extension of already existing buildings, emission of location maps for licensing acts or intervention on buildings, all require permission from the Municipality. In order to grant such permissions, the municipal services need to have access to updated information on the land. Such land use information is also required for municipal tax administration. Therefore, the basis for all these activities depends heavily on the spatial information in the form of maps. Such detailed maps should be updated regularly.

The current framework of the Portuguese large scale map production is very time-consuming and expensive, since it is based on manual editing of orthophotos to comply with very demanding technical specifications. Consequently, the time lag between availability of updated maps is generally 10 years.

To overcome the data update periodicity, we propose a methodology based on VHR imagery and Light Detection And Ranging (LiDAR) data to detect and identify changes in an out-dated cartography. The aim of this analysis is to highlight those areas where changes most likely have occurred. This new product, in a first step, can be analyzed by municipal technicians that: (1) will decide, based on analysis of the image and related information, if the marked spot is in fact a change area (a new urbanization or a built-up object that was demolished), and if so, (2) digitize the new buildings into the old cartography, or eliminate it, thus producing an updated map, if the *alarm layer* is accurate enough, or send a topographer to collect it directly. If the spot is considered to be a false detection, than the technician will (3) eliminate it. All these actions can be done using a GIS interface where the most recent building map is overlaid on the high-spatial resolution imagery, and where attributes like ongoing urban processes or the Constraints Master Plan map, can also be seen.

[1] teresasantos@fcsh.unl.pt

Such methodology can be used by the municipality to keep its cartographic database of urban areas up-to-date between two official map products, with higher thematic and positional accuracy.

1. Study area and dataset

The cartographic update was tested in a study area located in the oriental part of the city of Lisbon. The selected area occupies 64 ha (800 m × 800 m) (Figure 1), and is characterized by a diverse land cover that includes herbaceous vegetation, lawns, trees and agricultural plots, bare soil, single and multi-family housing, a school, industrial properties, and roads and rail networks.

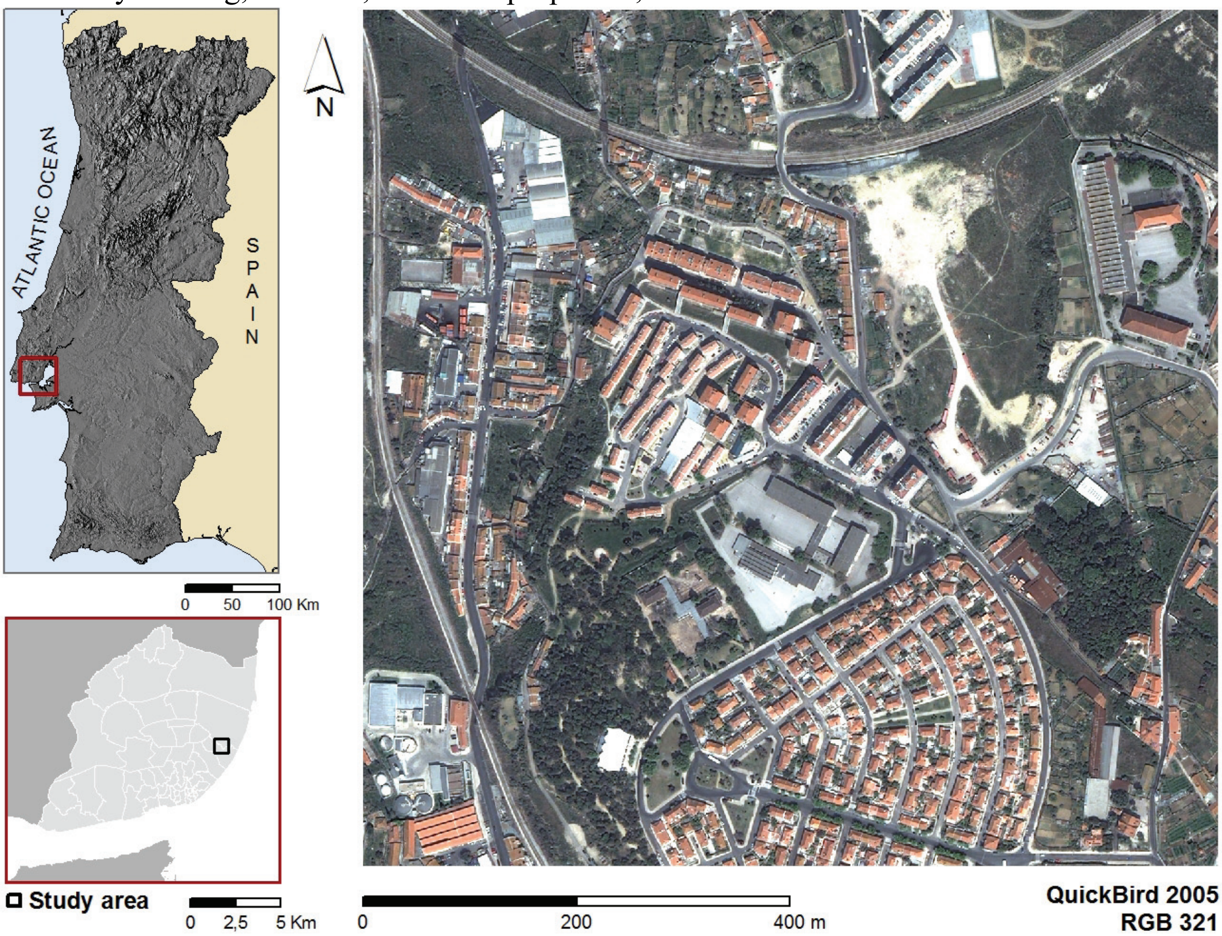

Figure 1. Study area location in the city of Lisbon.

The database includes planimetric, spectral, altimetric data. The map to be updated is the Lisbon's Municipal Cartography (Carto98), available at the Lisbon City Hall. This cartography was produced in a vector format, for the year 1998, and has a scale of 1:1000.

The spectral data is a QuickBird image acquired in April 14, 2005. The image has a spatial resolution of 2.4 m in the multispectral mode (visible and near-infrared bands), a pixel size of 0.6 m in the panchromatic mode, and a radiometric resolution of 11 bits. The image used in this study has an off-Nadir angle of 12.2º.

Furthermore, two altimetric datasets were also explored in this study. One set is derived from a LiDAR point cloud, and the other is derived from cartography. From a LiDAR flight performed in 2006, a surface image was produced based on the second return, with 1 m resolution. This image represents the Digital Surface Model (DSM) of the study area. From 1:1000 scale cartography of 1998, a set of elevation mass points and contours were retrieved.

2. Methodology

2.1. Pre-processing

In this stage, the altimetric data was processed in order to obtain the Digital Terrain Model (DTM). The DTM was then used to orthorectify the QuickBird image and to derive the normalized Digital Surface Model (nDSM). All files were geometrically corrected to attribute a common coordinate system (PT-TM06/ETRS89).

The DTM was obtained from the elevation mass points and contours. Firstly, a Triangulated Irregular Network (TIN) was generated and then converted to a grid with 1 m of resolution. This final file corresponds to the DTM for the study area.

The nDSM was then produced based on the DTM and the DSM. The nDSM was obtained by subtracting the DTM to the DSM image. This raster file stores the height of all elements above and below the terrain (Figure 2).

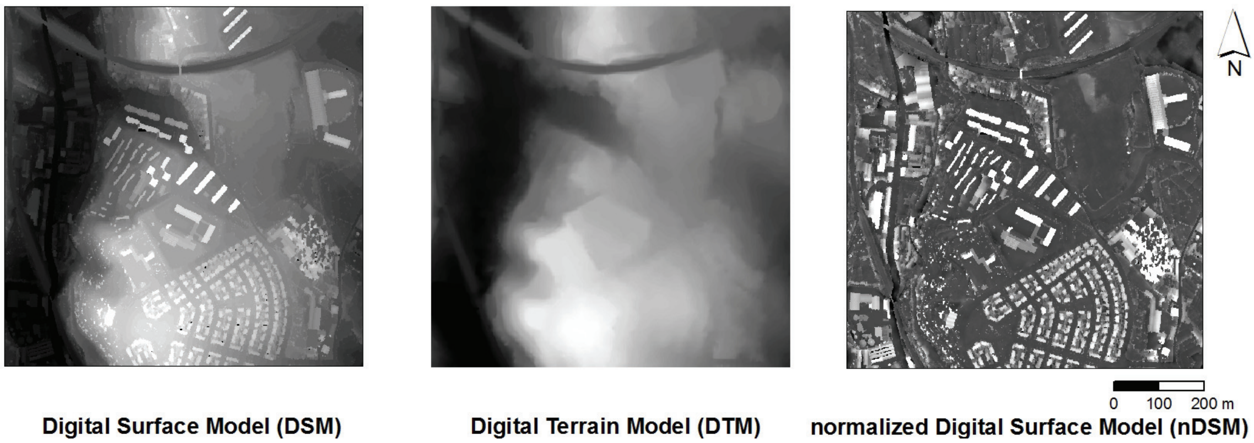

Figure 2. Altimetric data used in the study area.

The QuickBird imagery was orthorectified in order to reduce the geometric distortions introduced by the terrain. Previously, a pansharp image of the multispectral and panchromatic bands was produced for visual benefit. The orthorectification of the multispectral, pan, and pansharp images was performed based on the Rational Polynomial Coefficients (RPCs), available with the image, and a set of 36 ground control points retrieved from the 1:1000 planimetric and altimetric cartography of 1998, available from the Lisbon City Hall. For the elevation reference, the DTM, produced as described above, was applied. A 2^{nd} order polynomial was selected for the transformation. To evaluate the process, 40 check points, well distributed across the image, were used. The obtained Root Mean Square Error (RMSE) was less than one pixel. Afterwards, a Normalized Difference Vegetation Index image (NDVI) was produced to integrate the dataset for feature extraction.

The nDSM and the planimetry were well registered. However, the QuickBird's off-Nadir angle of 12.2º was responsible for the not so accurate registration between the imagery and the altimetric and planimetric dataset.

2.2. Land Cover Classification for 2005/06

The first step for updating the 1998 Cartography was to classify the QuickBird image from 2005, using the 2006 nDSM surface for better discrimination of the objects of interest. The goal is to produce a map with the land cover of 2005/06 (LCM).

The class of interest for map updating is the built up class. Nevertheless, other land cover classes are also considered in the classification system, in order to allow a good extraction. Regarding the building's rooftops, three classes were considered. One for the red tile cover, the most common material in the study area. Another class for the brighter roof material like light tin. And a class for remaining materials, that include dark tin, dark tile and fibrocement. Regarding the paved areas, three classes were identified: railways, roads and other impermeable surfaces that include streets, sidewalks and other paved material. Although present in the study area, a class for bare soil was not considered in this classification schema, since it did not affect the extraction of the built-up classes.

The nomenclature is organized in three levels of detail. The 1st level includes the classes "Urban" and "Vegetation". On the 2nd level, three classes are defined: "Building", "Pavement", and "Vegetation". On the 3rd level, 7 classes are identified: "Building with red tile roof", "Building with bright roofs" and, "Building with other roof", "Road", "Railway" and "Other impermeable surfaces", and "Vegetation" (Table 1).

Table 1. Land cover nomenclature for 2005/06

Land cover classes in 2005/06		
Level 1	Level 2	Level 3
Urban	Building	Building with red tile roof
		Building with bright roof
		Building with other roof
	Pavement	Railway
		Road
		Other impermeable surfaces
Vegetation	Vegetation	Vegetation

All feature extraction was performed in Feature Analyst 4.2 (by VLS) for ArcGIS (ESRI). The classification is based on a supervised approach. The first step is the manual digitizing of training areas for each class, followed by the definition of parameters like the number of bands to be classified, the type of input representation, and aggregation.

For the extraction, several data sets were used simultaneously as input: pansharp and multispectral QuickBird imagery, the NDVI grid, and the nDSM layer. The pansharp QuickBird image is of fundamental importance because it is the main reflectance layer and determines the scale and resolution (spatial detail) that features can be extracted. The approach selected in this work used, in a first stage, the possibility of classifying the study area in two major classes - "Vegetation", "Urban" – and in the subsequent stages the level 3 classes were extracted independently. The parameters that produced the best extraction results for each element type are available in [1].

The feature extraction stage was difficult due to the complex morphology and to the spatial heterogeneity of the study area [2]. Several iterations took place after the initial training in order to obtain the final classes. Such operations included removing clutter and adding missing data to allow the classifier to learn and produce a better extraction.

After feature generalization, an accuracy assessment of the quality of the LCM was conducted. Since the goal is to update the built-up classes in the 1998 Cartography, the quality assessment of the land cover extracted from the imagery dataset was applied only to the LCM level 2 class "Building". To evaluate the quality of spatial information automatically extracted from images, based on the concept of reference value, it is necessary to measure levels of compliance with information from an independent source. This reference data was a map obtained by visual interpretation of the same source data. All the discernible features belonging to the class of interest were digitized, without limits of size or shape. To assess the overall thematic quality of building extraction, the spatial overlap between classified and reference data is used.

The analysis indicated an Overall Accuracy of 73%. A 6% Commission Error was obtained. The Omission Error was 24% and occurred mainly in places where roofs were in the shadow, or were in different states of conservation, or where elevator shafts were present.

2.3. Map Updating

Map updating begins by selecting the classes of interest from the 1998 Cartography – "Buildings", "Annexes" and "Shacks" (BAS) – and then evaluate their status based on two datasets: the nDSM from 2006 and the Land Cover Map from 2005/06 (LCM). The nDSM is used to characterize the average height of every element above ground. From the LCM, only three classes are used: the level 1 classes "Urban" and "Vegetation", and the level 2 class "Building". This option is due to the fact that the remaining level 2 class "Pavement", only achieved a moderate overall accuracy of 65% [1], making it not reliable enough for map updating.

The update is based on a change\no-change approach. Since we are analysing built up classes, only three classes are possible in 2005/06: "No-Change" (the class in 2005/06 is the same as in 1998), "Change to Vegetation" (removed features) or "Change to New Building" (built-up features):

- If an object is labelled as "No-Change", then its class and geometry are the same as in the original 1998 Cartography.
- If an object is labelled as "Change", then two classes are possible, "Vegetation" or "New Building". The 1998 built up objects identified as "Vegetation" in the LCM, maintain their geometry and receive a new classification. The objects identified as "New Building", have their geometry based on the LCM.

The update is done through map algebra operations over the available datasets, and follows four hierarchic rules:

- Rule 1 – every object identified as "Buildings", "Annexes" or "Shacks" in 1998, and having height above 3 m in 2006, is labelled as No-Change T1 (FromBAS ToBAS);
- Rule 2 – every object identified as "Buildings", "Annexes" or "Shacks" in 1998, shorter or equal to 3 m in 2006, and classified as "Urban" in LCM, is labelled as No-Change T2 (FromBAS ToBAS);
- Rule 3 – every object identified as "Buildings", "Annexes" or "Shacks" in 1998, but classified as "Vegetation" in LCM, is labelled as Change T1 (FromBAS ToVegetation);
- Rule 4 – every object not identified as "Buildings", "Annexes" or "Shacks" in 1998, but greater than 1 m in 2006, and classified as "Buildings" in LCM, is labelled as Change T2 (ToNewBuilding);

The final map of the study area, updated for 2005/06 (MapUp) according to the previous schema, has three classes: "No-Change", "Change to Vegetation" and "Change to New Building" (Figure 3).

In the period under analysis – 1998 to 2005/06 – the main types of change identified in the study area were: shacks' eradication and building demolitions (industrial properties), newly built industrial sites (e.g., the wastewater treatment plant, located in the bottom left corner of the map), as well as new residential housing (e.g., two multi-family houses).

2.4. Quality of Detection of Building Change

After map updating, its quality was evaluated based on visual interpretation. A census, instead of sampling approach, was conducted in order to fill the error matrix. Two analyses were performed: one for all the objects with area greater than 20 m^2, and another for all the objects with an area greater than 10 m^2 and less or equal to 20 m^2. With this segmentation, we intend to evaluate the impact of smaller features in the map's quality. Objects having 10 m^2 or smaller were not evaluated since many of them are errors introduced by the misregistration between the QuickBird imagery and the nDSM and 1998 Cartography.

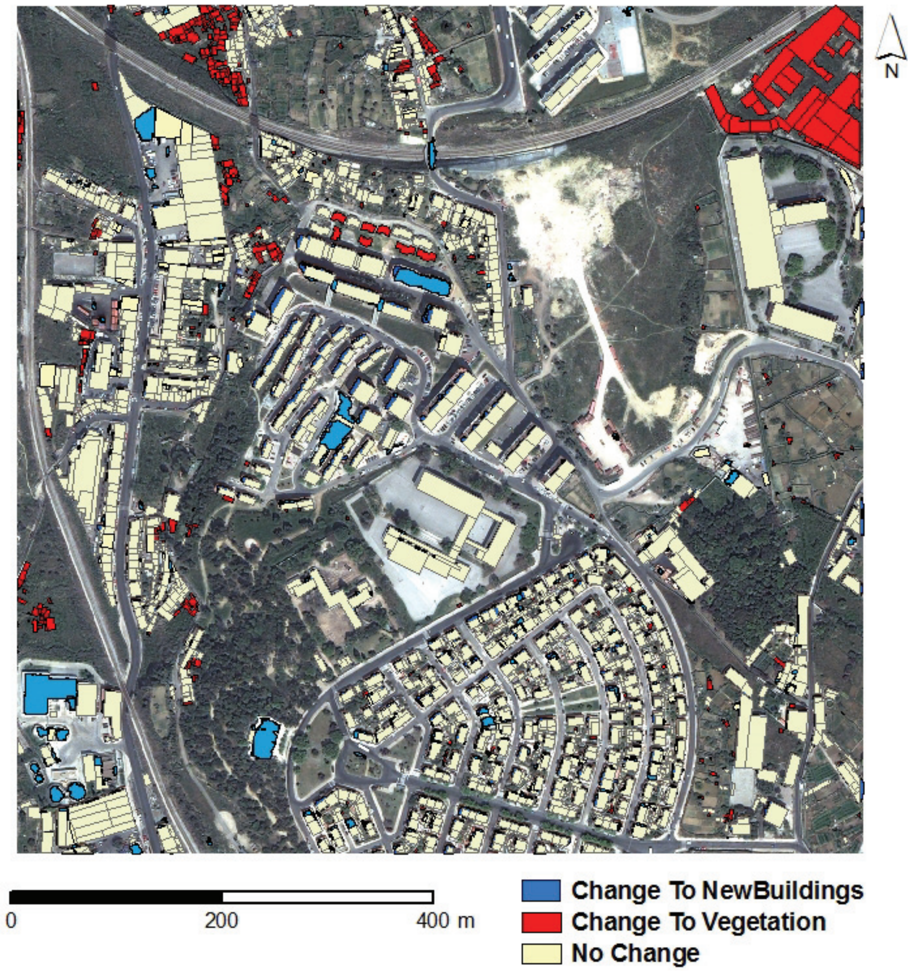

Figure 3. Map updated to 2005/2006 (MapUp).

Among all objects with an area larger than 20 m^2 that were visually inspected (1239 objects), 9 were excluded from the validation because the final class could not be confirmed with a sufficient level of confidence. From this analysis, we can confirm that the class "Change To Vegetation" was well mapped. Only 3 features that were under trees were incorrectly classified as change. The "Change To New Building" was also very accurate, with no mistakes. The class "No-Change" was also well mapped, but had 15 misclassified features. From these, 13 objects were in fact demolished buildings that changed to road/pavement in 2005/06. For further analysis the results of the validation were grouped into two classes – "Change" and "No-Change" - and an error matrix was populated (Figure 4). Several indices were calculated to assess the quality of the objects larger than 20 m^2. From this analysis, we conclude that the updated map has an Overall Accuracy of 99% and a KHAT statistic of 95%.

Map \ Reference	Change (objects)	No-Change (objects)	Total row
Change	198	3	201
No-Change	15	1014	1029
Total column	213	1017	1230

Figure 4. Error matrix for the Change and No-Change objects with areas larger than 20 m^2

The same evaluation was made for those objects with area larger than 10 m^2 and smaller or equal to 20 m^2. From the total of 688 objects analysed, 30 were rejected from the validation because the final land cover class could not be visually confirmed. From the 116 objects classified as

"Change To Vegetation", 14 were false changes located under trees. The accuracy of class "Change To New Building" was not very satisfactory, with 41 errors among the 69 validated objects. The "No-Change" class was well mapped, but several errors occurred, once more, in features that changed to pavement. From the error matrix, the thematic accuracy indices were calculated (Figure 5). We conclude that the quality of the update for small objects is lower than those larger than 20 m^2, but still very high, with an Overall Accuracy of 89% and a KHAT statistic of 70%.

Reference / Map	Change (objects)	No-Change (objects)	Total row
Change	130	55	185
No-Change	19	454	473
Total column	149	509	584

Figure 5. Error matrix for the Change and No-Change objects with areas larger than 10 m^2 and smaller or equal to 20 m^2

3. Conclusions

The proposed methodology is based rather on a semi-automatic verification and update of existing large-scale cartography, available in a GIS format, using recent image data as reference information. The advantage of the developed system is that it reduces the manual efforts of a human operator, saving time and, probably, costs, allowing concentrating attention only on those changed areas, and consequently, making it a more efficient process. Furthermore, the semi-automatic system contributes to a more effective decision-making process, through the production of new information in a regular basis. The developed change detection method enables to: identify areas of land cover change and, indicate the type/ direction of change.

In this extracting experience with VHR imagery, several problems were detected. Complex shapes and multilevel structures could not be well identified. Differentiating between building extensions was a difficult task to perform in this semi-automatic environment. The presence of different roof covers and structures in the same building, also contributed to a poor identification and delineation of those features.

For municipal planning, and according to the technical specifications of large-scale cartography (1:5000 scale and higher), map production based on VHR images and photogrammetry is still necessary to guarantee that each uniquely identified feature is well delineated and stored in the database as a geometric entity together with a list of attributes. Nevertheless, all land information requires periodical updates. This task involves searching for small areas of change within an image. Using digital photogrammetric workstations and extremely high resolution aerial photography for this change detection process takes up valuable resources [3]. In this context, semi-automatic classification of VHR images can give a great contribution by speeding up the change detection process, and alert the technicians for those areas where potential land change has occurred, thus combining automatic change detection and human-computer interaction.

Acknowledgements

This work was performed under the framework of project GeoSat - Methodologies to extract large scale GEOgraphical information from very high resolution SATellite images, funded by the Foundation for Science and Technology (PTDC/GEO/64826/2006). The authors would like to thank Logica for the opportunity of using the LiDAR data set.

References

[1] T. Santos, S. Freire, A. Navarro, F. Soares, J. Dinis, N. Afonso, A. Fonseca, J. A. Tenedório. Extracting buildings in the city of Lisbon using QuickBird images and LIDAR data. *Proceedings of GEOBIA 2010 - GEOgraphic Object-Based Image Analysis*, 29 June - 2 July, Ghent, Belgium.

[2] S. Freire, T. Santos, N. Gomes, A. Fonseca, J. A. Tenedório. Extraction of buildings from QuickBird imagery for municipal use – the relevance of urban context and heterogeneity. *Proceedings of EARSeL 2010*, 29 May - 2 June, Paris, France.

[3] D. A. Holland, D. S. Boyd, P. Marshall. Updating topographic mapping in Great Britain using imagery from high-resolution satellite sensors, *ISPRS Journal of Photogrammetry & Remote Sensing*, **60** (2006), 212–223.