EXTRACTION OF BUILDINGS FROM QUICKBIRD IMAGERY FOR MUNICIPAL PLANNING PURPOSES: QUALITY ASSESSMENT CONSIDERING EXISTING MAPPING STANDARDS

S. Freire [a,*], T. Santos[a], A. Navarro[b], F. Soares[b], J. Dinis[b], N. Afonso[c], A. Fonseca[c] , J. A. Tenedório[a]

[a] e-GEO, Faculdade de Ciências Sociais e Humanas, FCSH, Universidade Nova de Lisboa, Portugal - (sfreire, teresasantos, ja.tenedorio)@fcsh.unl.pt

[b] University of Lisbon, Faculty of Sciences, LATTEX-IDL, Portugal - (acferreira, fsoares)@fc.ul.pt, dinis.joel@gmail.com

[c] National Laboratory for Civil Engineering (LNEC), Lisbon, Portugal - (nafonso, anafonseca)@lnec.pt

KEY WORDS: QuickBird, Feature extraction, Buildings, Urban planning, Quality assessment, Lisbon

ABSTRACT:

Many municipal activities require updated large-scale maps that include both topographic and thematic information. For this purpose, the efficient use of very high spatial resolution (VHR) satellite imagery suggests the development of approaches that enable a timely discrimination, counting and delineation of urban elements according to legal technical specifications and quality standards. Therefore, the nature of this data source and expanding range of applications calls for additional methods and metrics to assess the quality of the extracted information which go beyond traditional thematic accuracy alone. The present work concerns the feasibility of VHR satellite imagery as an alternative source of geospatial information for large scale mapping to assist urban planning in Portugal. Feature extraction software was employed to map buildings present in a pansharpened QuickBird image of Lisbon. Quality assessment was exhaustive and involved comparisons of extracted features against a reference dataset, introducing cartographic constraints from scales 1:1000, 1:5000, and 1:10000. The spatial data quality elements subject to evaluation were: thematic (attribute) accuracy, completeness, and geometric quality based on planimetric deviation from the reference. Tests were developed and metrics analyzed considering thresholds and standards for the different mapping scales most used by municipalities. Results show that values for completeness varied with mapping scales and results were only slightly better for scale 1:10000. Concerning the geometric quality, a large percentage of extracted features met the strict topographic standards of planimetric deviation for scale 1:10000, while no buildings were compliant with the specification for scale 1:1000.

1. INTRODUCTION

A spatial component is associated with the majority of municipal activities, namely in urban planning and management. At this level, decision-making is supported by large-scale spatial data that include both topographic and thematic information, but which frequently become outdated due to the dynamics of the urban environment. Very high spatial resolution (VHR) satellite imagery can be used for faster updating of municipal spatial databases, but the nature of this recent data source and target features, the object-based image analysis, and expanding range of applications call for additional methods and metrics to assess the quality of the extracted spatial information (see Zhan et al., 2005). These go beyond traditional pixel-based thematic accuracy alone, requiring the assessment of discrimination/classification, detection/counting, and delineation of features of interest. Also, an efficient operational use of VHR satellite imagery suggests the development of approaches for mapping urban features that conforms to legal technical specifications and quality standards. Meeting all these requirements may change quality assessment from a generic process to one specific for each feature type.

Buildings are a major urban element and one of the main feature classes of interest for a municipality, whose 'correct' semi-automatic extraction from imagery remains a challenging task. To obtain a cartographic product, most of the challenge results from the interplay of several factors, namely the object and its context, nature of imagery, and mapping requirements and constraints. Despite the many methodologies proposed for feature extraction, none has so far proved to be effective in all conditions and for all types of data (Salah et al., 2009).

For the image analyst / map producer the challenge may be limited to handling the necessary stages of image pre-processing, image segmentation, and generalization of features to produce a map. In the present context, the very quality assessment of extracted buildings is a complex endeavour for which there is no consensual or standard approach.

Van Coillie *et al*. (2008) presented a methodology for a supervised, objective evaluation of segmentation quality based on quantitative similarity measures. The methods were tested on a single house, and its manual digitizing used as reference. Eight quality measures were tested to compare different segmentations layers with the reference one. The discrepancy quality measures included the number of segments that have their centroid in the reference polygon, difference in total area and total perimeter, difference in shape complexity, average distance between edge pixels and cumulative distance from the reference.

* Corresponding author.

Regarding polygon generalization, despite being the subject of significant research, there is still a need for comprehensive investigation (Podolskaya et al., 2007). Khoshelham et al. (2009) have conducted a detailed comparative analysis of five automated methods for building detection, but used pixel-based metrics for accuracy assessment.

Accuracy assessment of thematic maps, based on map comparison, has often neglected cartometric quantities (Dungan, 2006). In the case of buildings, topographic maps represent the building footprint according to scale-dependent constraints. Vu et al. (2009) propose a multi-scale solution based on mathematical morphology for building extraction using LiDAR and image data. This approach allows extraction of complex buildings as scale-dependent multi-part objects and to capture building footprint.

Introducing mapping specifications in quality assessment of extracted features, Gianinetto (2008) tested roads and buildings extracted from pansharpened QuickBird imagery for updating large-scale topographic databases of urban areas. Results showed that updating of 1:10000 scale was always compliant with standards, while updating of scale 1:5000 was only possible in certain situations. However, test features were manually digitized.

The GeoSat research project, which involves the Lisbon City Hall, aims at developing methods to expedite the production of geographic information for municipal planning and land monitoring, and investigates the potential of VHR satellite imagery and Geographic Object-Based Image Analysis (GEOBIA) (Hay and Castilla, 2008) for detection and mapping of urban features and their integration into operational urban planning and management activities. Previous work (Santos et al., 2009) has explored and proposed detailed vector-based metrics for accuracy assessment of QuickBird-derived buildings, but without taking map standards into account.

The goal of the present research is to introduce existing scale-based mapping constraints from official specifications in the process of quality assessment of features extracted from QuickBird imagery. The motivation is to evaluate the feasibility of features extracted from Very-High Resolution (VHR) imagery by semi-automatic methods to readily integrate a municipal GIS database. First, buildings were extracted from the image using feature extraction software and ancillary data. The second part included the development and testing of quality assessment procedures considering thresholds and standards for the different mapping scales used by municipalities, analysis of metrics, and discussion of results.

2. STUDY AREA AND DATA

2.1 Study area

For this study, an area located to the northeast of the downtown of the city of Lisbon, Portugal, was selected (Figure 1). This area occupies 64 ha (800 m by 800 m), and has a diverse land use/land cover (LULC) that varies from urban to open field with and without vegetation. It includes trees, lawns, herbaceous vegetation and agricultural plots, bare soil, a school, industrial properties, roads and rail networks, and residential housing. This latter use includes a mixture of single homes and multi-story apartment buildings.

Figure 1. Study area in the city of Lisbon and pansharpened QuickBird image

In the study area 627 building blocks were identified, having a wide variety of roof types. Red tile roofs are the standard for residential buildings, and buildings with this coverage comprise 53% of the total identified.

2.2 Data sets

Several spectral, altimetric, and planimetric spatial data sets were used for feature extraction and quality assessment, namely: a pansharpened QuickBird image, its multispectral bands, a derived vegetation index from these bands (NDVI), a normalized Digital Surface Model (nDSM), and a vector reference map of building polygons (Table 1).

Data set	Data type	Resolution (m)
QuickBird pansharp	Raster	0.6
QuickBird multispectral	Raster	2.4
NDVI	Raster	2.4
nDSM	Raster	1
Ref. map of bldgs.	Vector	-

Table 1. Main features of data sets used

The QuickBird imagery was acquired in April 14, 2005 with an off-Nadir angle of 12.2º. 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 imagery was pansharpened to the spatial resolution of the panchromatic band in PCI Geomatica, and orthorrectified in order to reduce the geometric distortions introduced by the relief and to attribute a national projected coordinate system (ETRS89-PT-TM06). The orthorrectification was performed using the Rational Polynomial Coefficients (RPCs) provided with the image, and a set of 36 ground control points retrieved from the 1:1000 planimetric and altimetric cartography of 1998. The RMSE (Root Mean Square Error) of the transformation was less than one pixel.

3. METHODOLOGY

Pre-processing of data for the present study has included orthorrectification and pansharpening of imagery, computing the NDVI, and production of the nDSM grid. For more details on this stage see Santos *et al.*, 2010.

3.1 Feature extraction

Extraction of features (polygons) from the imagery was performed using Feature Analyst 4.2 (VLS), as an extension for ArcGIS (ESRI). Feature Analyst (FA) is a GEOBIA application that conducts an internal "hidden" segmentation of the image that allows to classify and extract only those features belonging to the class of interest.

The classification is based on a supervised approach, so the initial step is the identification of training samples for each class, followed by the definition of parameters such as the number of bands to be classified, the type of input representation, and level of aggregation. The classifier uses feature characteristics such as spectral response/color, size, shape, texture, pattern, shadow, and spatial association, for feature classification. After an initial classification, there is the possibility to remove clutter or add missing areas. This hierarchical learning adaptive process allows to iteratively improve the image classification. The classified map can then be post-processed to aggregate and geometrically generalize the features.

For the extraction of buildings with red tile roofs, several data sets were used simultaneously as input: pansharpened and multispectral QuickBird imagery, the NDVI grid, and the nDSM layer. The pansharpened 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 best extraction result for this class was obtained after two iterations of removing clutter, using 24 training areas and the following parameters: 'Manhattan 5' for the input representation, masking out vegetation and white roofs, and aggregation of 100 pixels.

Feature Analyst includes post-processing tools that allow geometric generalization of polygons (smoothing, square up), especially important for man-made features such as buildings and considering the current purpose for their extraction. The extracted 'raw' buildings with red tile roofs were generalized (squared up) with the following parameters: 1 m smoothing tolerance, 6 pixel squaring tolerance, considering adjoining features and all likely orientations.

3.2 Quality assessment

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 can be obtained from a field survey (e.g. GPS collection), from an existing map having acceptable accuracy, or from a map obtained by visual interpretation of the same source data.

For the study area there is no official 1:10 000 scale map, and the off-nadir look of the QuickBird image used and the resulting 'leaning' of buildings would prevent a fair comparison with an official planimetric map if it existed. Also, topographic maps represent the building footprint, whereas satellite or aerial imagery capture its roof. Due to these limitations, a reference map of building blocks was created by visual analysis and manual digitizing over the pansharpened image. All the discernible features belonging to the class of interest were digitized, without limits of size or shape.

The quantitative analysis of quality was exhaustive and took place in two stages, using ArcGIS 9.3 (ESRI). The first stage involves the comprehensive analysis of the entire class for classification error, and results in the evaluation of its thematic accuracy and completeness (lack of errors of omission and commission). The second stage occurs only for those class features that represent the same object in the reference and classification sets (1-1 relation) and assesses geometric quality and integrity based on tolerances for three mapping scales: 1:1000, 1:5 000 and 1:10 000.

The ambitious rationale was to test the methodology considering scales for maps frequently used at the municipal level, although 1:5000 and 1:1000 are beyond the mapping scales suitable for the pansharpened QuickBird image. Its 0.6-m pixel size sets 1:6000 as the limit for largest mapping scale possible (see Steiniger, 2008). This problem is limited by the fact that classification and reference sets are produced from the same image.

The selected scales imply strict cartographic constraints from technical specifications adopted and published by the Portuguese Geographic Institute for digital topographic data (Table 2).

Scales	Area (m^2)	Tolerance (m)	
		RMSE	90%
1:1 000	4	0.18	0.27
1:5 000	4	0.75	1.25
1:10 000	20	1.50	2.30

Table 2. Constraints for size and planimetric tolerance of features for selected scales of digital topographic data

For representation at 1:1000 scale, a feature should be larger than 4 m^2, and planimetric deviation cannot be greater than 0.18 m (RMSE) or has to be smaller than 0.27 for 90% of the sample.

Thematic quality and completeness were assessed with application of only the area constraint, while geometric quality was evaluated with enforcement of both the area and planimetric tolerance constraints.

Imposing the area constraint on the extraction and reference data sets resulted in no change in the number of features for the larger scales, since the smallest building block is larger than the 4 m^2 threshold (see Table 3). For scale 1:10 000 the number of features considered for analysis decreased from 330 to 313 in the reference, and from 316 to 272 in the generalized extraction.

To assess the overall thematic quality of building extraction, the spatial overlap between classified and reference data is used. This area-based test essentially evaluates the accuracy of the classification in terms of its extent and spatial distribution. The two vector sets are overlayed (union), and the overall thematic accuracy (TA) is obtained by dividing the area common to the two sets (intersection) by the area of union, according to Equation 1:

$$TA = \frac{A \cap B}{A \cup B} \qquad (1)$$

The analysis of completeness is feature-based and made using a reciprocal approach involving the features' centroids: first, the extracted polygons that contain centroids of the reference buildings are selected; those which were not selected have no correspondence in the reference and stand for error of commission. Then, the centroids of the selected extracted features are used to select reference polygons that contain them; the reference polygons not selected have no correspondence in the extraction and stand for the error of omission.

The assessment of geometric quality is based on the rationale that polygon area and shape are determined by its outline (edges), so it makes sense to analyze the latter for deviation from a reference ('ground truth'). Evaluation was initiated by selecting features having 1-1 cardinality among extracted and reference sets and excluding 1-n and n-1, using overlap.

Planimetric tolerance is a constraint devised for point-based testing that was adapted for verifying the compliance of polygons, by buffering each reference feature using the tolerance distances for each scale and calculating the percentage of the extracted building outline that falls inside the tolerance, i.e. is compliant. This process is illustrated in Figure 2.

Since compliance with a RMSE standard is difficult to verify exhaustively in this way, verification of the 90% criteria was adopted instead, with the full building outline being analyzed. If 90% of more of the outline meets the tolerance, the building is compliant.

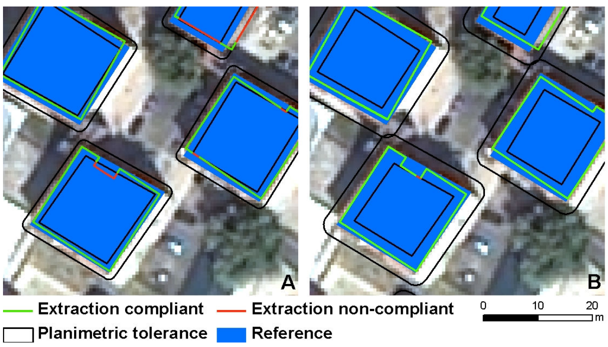

Figure 2. Detail of planimetric tolerance compliance test for scales 1:5 000 (A) and 1:10 000 (B)

Finally, to assist with exploration and explanation of previous results, a widely-used Shape Index (SI) was computed for extracted and reference buildings and for the different scale sets. The Shape Index was calculated according to Equation 2 and compares the perimeter of the element with the perimeter of a circle with the same area, therefore it is independent of scale. SI has a value of 1 for a circular region and increases as shapes become increasingly noncircular.

$$SI = \frac{Perimeter}{2.\sqrt{\Pi.Area}} \qquad (2)$$

4. RESULTS AND DISCUSSION

4.1 Feature extraction

The feature extraction stage was complicated due to the complexity of the study area and the heterogeneity of the buildings present, which also exists within the specific class of interest. Although individual buildings were used as training samples, FA has revealed to be unable to return adjacent buildings as separate features, even when a marked but narrow separation can be visually identified in the image. For this reason, quality assessment had to be conducted at the level of building block, these being equal to the building in the case of individual features.

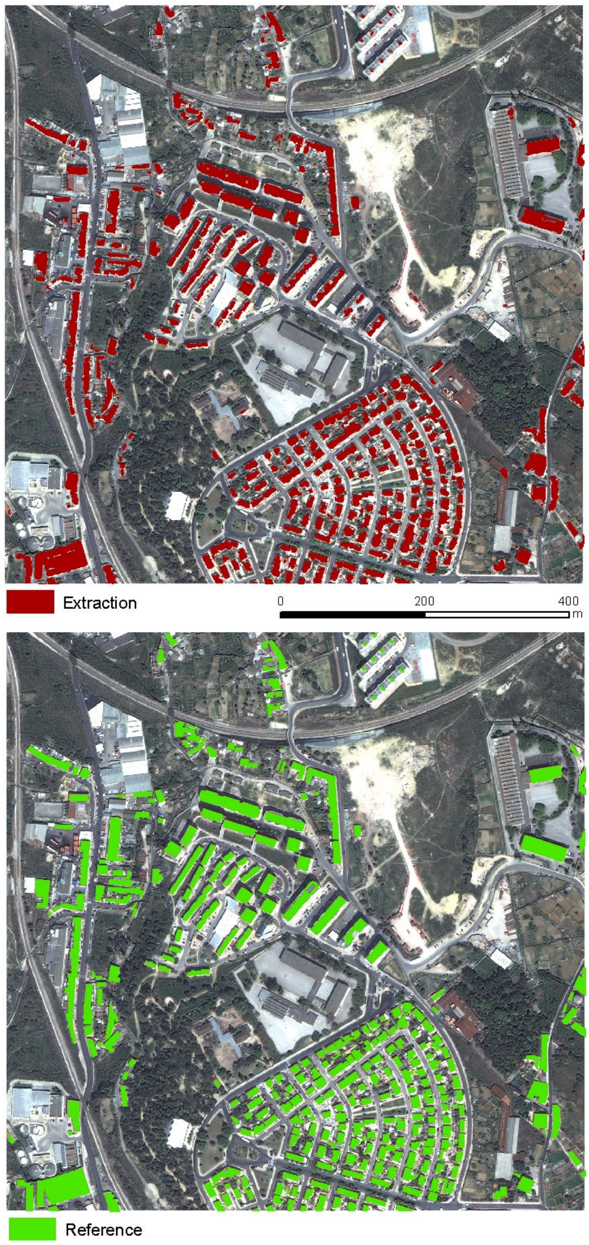

Figure 3. Extraction and reference data sets in study area

The best feature extraction result returned 317 'raw' buildings with red tile roofs, compared to the 330 mapped in the reference data set. The overall detection and distribution of buildings in these data sets can be compared in Figure. 3.

The extracted features have areas varying from 9.4 to 2359 m^2, and a mean size of 163.4 m^2. The size of the smallest feature detected is mostly related to the level of aggregation selected. The standard deviation of 201.8 gives some indication of the heterogeneity of building sizes, further confirmed by the reference map. Table 3 shows for each data set of buildings the number and area of features.

Data set	No. of feat.	Area (m^2)			
		Min	Max	Mean	Std.
Extraction	317	9.4	2359.1	163.4	201.8
Generalized	316	5.1	2370.5	154.9	202.5
Reference	330	6.6	2473.9	181.9	229.1

Table 3. Number of building features and their area in each data set

The generalization step decreased the overall number of features to 316, by merging irregular polygons which were very close. Also, this procedure has generally lowered the size of features, which decreased 8.5 m^2 on average, while their actual mean size (Reference) appears to be larger (181.9). Still, the generalization has significantly improved the overall geometric quality of features, as illustrated in Figure 4.

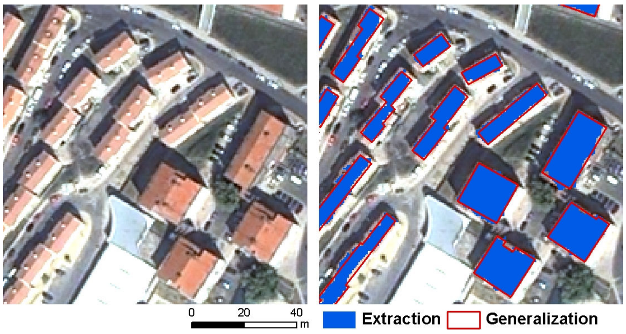

Figure 4. Detail of red tile roofs in study area and their extraction and generalization

4.2 Quality assessment

The overall thematic accuracy of building extraction was 72.2% for both 1:1000 / 1:5000 and 1:10000 map scales.

Regarding level of completeness, Omission error was higher than Commission and was higher for scale 1:10000 (Table 4), because a higher percentage of reference buildings larger than 20 m^2 were missed by the extraction. However, error of Commission was lower for the smaller scale, only 8% (22 polygons in 272 extracted).

Scales	Omission		Commission	
	No.	%	No.	%
1:1 000/1:5 000	68	20.7	45	14.2
1:10 000	71	22.7	22	8

Table 4. Completeness accuracy

Concerning geometric quality, 245 pairs of buildings were compared at the larger scales and 223 were tested for the scale 1:10000. Results were not satisfactory for scales 1:000 and 1:5000 (Table 5). For the largest scale, no features attained the 90% compliance value, while it was attained by only 26% of buildings at scale 1:5000, with a mean value of compliance of 78%. At scale 1:10000 89% of features are compliant, with a mean compliance of 96%. These results are in line with those obtained by Gianinetto (2008).

	Scales		
	1:1 000	1:5 000	1:10 000
Min	0	19.3	54.9
Max	78.2	100	100
Mean	20.2	78.1	96.1
Std.	12.7	16.1	8.8
No. of features	0	63	198
% of features	0	26	89

Table 5. Compliance values (%) for features for the 90% planimetric tolerance, by scale

Figure 5 illustrates distribution of building compliance in the study area for the two smaller scales. Such a reliability map, when available, can be used to guide manual editing of features to correct errors (Benz et al., 2004).

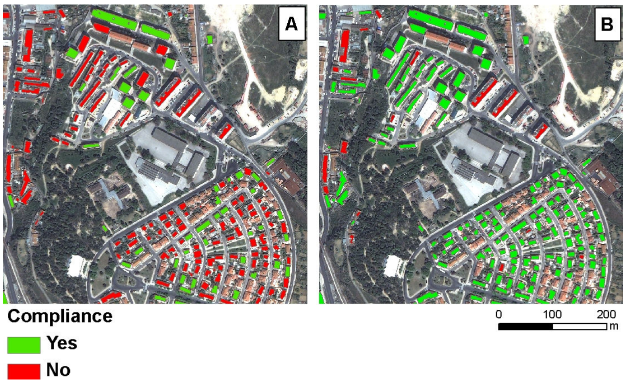

Figure 5. Compliance map for planimetric tolerance for scales 1:5 000 (A) and 1:10 000 (B)

Analysis of correlation has shown that compliance is not linearly correlated to feature size or their shape index.

However, grouping features by classes of area and plotting their respective accuracy has revealed a slightly parabolic-type curve where compliance is highest for medium-sized features, and lowest for those smallest and largest (Figure 6). This indicates that there might be a 'preferred' building size for correct extraction and/or that the smallest and largest buildings display features that complicate their accurate extraction.

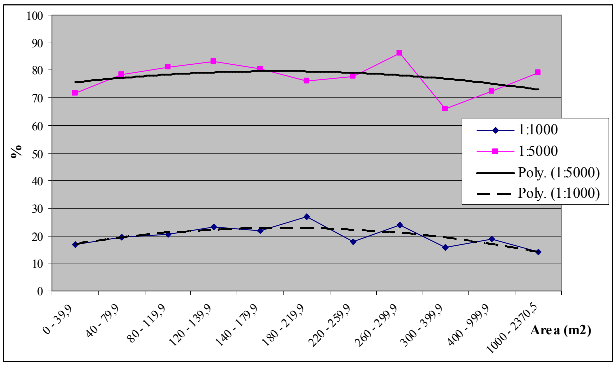

Figure 6. Average compliance of features by classes of area

5. CONCLUSIONS

The present work is an exploratory attempt to assess the quality of buildings extracted from VHR satellite imagery using semi-automated methods through analysis of similarity with a reference database. The main aim was to introduce mapping standards in a feature-based evaluation of different spatial quality elements.

Overall thematic accuracy was reasonable and invariant with introduction of area constraint for mapping scales. On the contrary, values for completeness varied with mapping scales and results were slightly better for scale 1:10000. However, testing spatially for completeness based on objects is a complex issue that needs improvement. The assessment of geometric quality and integrity revealed that strict topographic standards of planimetric deviation were only met at scale 1:10000, for a large percentage of extracted features. However, in order to produce an effective quality assessment tool, it would be important to integrate the different quality dimensions under one single metric that could be computed for each type of object.

Future developments also include assessing the quality of extracted natural elements (e.g., trees) and linear features (i.e., roads) that may involve conceptually different approaches. Additional cartographic standards should also be considered. The relevance of study area and heterogeneity in the feature extraction process will be the subject of future research.

REFERENCES

Benz, U.C., Hofmann, P., Willhauck, G., Lingenfelser, I., Heynen, M., 2004. Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GISready information. *ISPRS Journal of Photogrammetry & Remote Sensing*, 58, pp. 239–258.

Dungan, J.L., 2006. Focusing on feature-based differences in map comparison. *Journal of Geographical Systems*. 8, pp. 131-143.

Gianinetto, M., 2008. Updating Large Scale Topographic Databases in Italian Urban Areas with Submeter QuickBird Images. *International Journal of Navigation and Observation,* Vol. 2008.

Hay, G. J., Castilla, G. 2008. Geographic Object-Based Image Analysis (GEOBIA): A new name for a new discipline? In: *Object-Based Image Analysis – spatial concepts for knowledge-driven remote sensing applications*. Eds:T. Blaschke, S. Lang, G. J. Hay. Springer-Verlag., pp. 75-89.

Khoshelham, K., Nardinocchi, C., Frontoni, E., Mancini, A., Zingaretti, P., 2009. Performance evaluation of automated approaches to building detection in multi-source aerial data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 65(1), pp. 123-133.

Podolskaya, E.S., Anders, K.-H., Haunert, J.-H., Serester, M., 2007. Quality Assessment for Polygon Generalization. In: *5th Int. Symposium on Spatial Data Quality (SDQ 2007)*, Enschede, Netherlands, Vol. 34, Part XXX.

Salah, M., Trinder, J., Shaker, A., 2009. Evaluation of the Self-Organizing Map Classifier for Building Detection from LiDAR Data and Multispectral Aerial Images. *Journal of Spatial Science*, 54(2), pp. 1-17.

Santos T., Freire S., Navarro A., Soares F., Dinis J., Afonso N., Fonseca A., Tenedório, J. A., 2010. Extracting buildings in the city of Lisbon using QuickBird images and LiDAR data. In: GEOBIA2010 - *GEOgraphic Object Based Image Analysis, Ghent, Belgium.*

Santos, T., Freire, S., Boavida-Portugal, I., Fonseca, A. Tenedório, J.A., 2009. Accuracy assessment of features extracted from QuickBird imagery for urban management purposes. In: *33rd Int. Symposium on Remote Sensing of Environment*, Stresa, Italy.

Steiniger S., Castilla G., Hay G. J., 2008. From image-objects to maps: an assessment of cartographic requirements for GEOBIA. Presentation at GEOBIA 2008, Calgary, Canada, 7 August 2008.

Van Coillie, F., Pires, R., Van Camp, N., Gautama, S., 2008. Quantitative segmentation evaluation for large scale mapping purposes. In: *GEOBIA 2008 - Pixels, Objects, Intelligence. GEOgraphic Object Based Image Analysis for the 21st Century*, Calgary, Canada.

Vu, T.T., Matsuoka, M., Yamazaki, F., 2009. Multi-scale solution for building extraction from LiDAR and image data. International *Journal of Applied Earth Observation and Geoinformation*, 11(4), pp. 281-289.

Zhan, Q., Molenaar, M., Tempfli, K., Shi, W., 2005. Quality assessment for geo-spatial objects derived from remotely sensed data. *International Journal of Remote Sensing*, 26(14), pp. 2953 – 2974

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 Portuguese 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.