

# Multilevel Object-Oriented Classification of Quickbird Images for Urban Population Estimates

Claudia M. Almeida  
DSR-INPE, Brazil  
Av. Astronautas, 1758  
12227-010 São J. Campos  
Phone: +55-12-3945-6428  
almeida@dsr.inpe.br

Iris M. Souza  
DSR-INPE, Brazil  
Av. Astronautas, 1758  
12227-010 São J. Campos  
Phone: +55-12-3945-6426  
iris@dsr.inpe.br

Claudia Durand Alves  
DSR-INPE, Brazil  
Av. Astronautas, 1758  
12227-010 São J. Campos  
Phone: +55-12-3945-6434  
durand@dsr.inpe.br

Carolina M. D. Pinho  
DPI-INPE, Brazil  
Av. Astronautas, 1758  
12227-010 São J. Campos  
Phone: +55-12-3945-6471  
carolina@dpi.inpe.br

Madalena N. Pereira  
DSR-INPE, Brazil  
Av. Astronautas, 1758  
12227-010 São J. Campos  
Phone: +55-12-3945-6434  
madalena@dsr.inpe.br

Raul Q. Feitosa  
Catholic University of Rio  
R. Marquês São Vicente, 225  
22453-900 Rio de Janeiro  
Phone: +55-21-3527-1212  
raul@ele.puc-rio.br

## ABSTRACT

This paper is committed to explore object-oriented methods for the classification of Quickbird images, aiming to support future urban population estimates. The study area concerns the southern sector of São José dos Campos city, located in the State of São Paulo, Brazil. By means of a multi-resolution segmentation approach and a six-layer hierarchical classification network, homogeneous residential areas were identified in terms of density of occupation and building standards (single dwelling units or high-rise buildings). The classification network was built upon spectral, geometrical and topological features of the objects in each level of segmentation as well as upon their contextual and semantic interrelationships in-between the hierarchical levels. The final classification of homogeneous residential units was subject to validation, using an object-based Kappa statistics.

## Categories and Subject Descriptors

I.2.4 [Computing Methodologies]: Knowledge Representation Formalisms and Methods – *representations (procedural and rule-based), semantic networks.*

## General Terms

Experimentation, Design, Algorithms.

## Keywords

Object-oriented analysis, high resolution images, knowledge-based methods, semantic networks, population estimates.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

ACM GIS'07, November 7-9, 2007, Seattle, WA.

Copyright 2007 ACM 978-1-59593-914-2/07/11...\$5.00

## 1. INTRODUCTION

Works dealing with satellite imagery for population estimates have become more and more often in the latest years. In Northern Leicestershire, UK, a population density surface was conceived for this end [1]. Initially, a LS-TM image was registered onto the same geographical coordinates as the census data, and then, a supervised maximum likelihood classification was conducted using principal components analysis (PCA) as input, in order to identify pixels that could be classified as residential. The population figures were allocated to the residential pixels falling within each ward, through a simple division between population and number of pixels, so as to produce a crude dasymetric map of the population-per-pixel. And finally, a surface transformation was applied on the image using kernel areas of 0.25 km<sup>2</sup> and 1.0 km<sup>2</sup>, and the results could then be visualized through a fish-net representation, onto which a shade sequence has been superimposed according to the degrees of population density.

The use of texture and contextual measures to be associated with housing densities by means of linear regression models has also been proposed [2]. The housing density was considered as the dependent variable, and the mean reflectance values on the image bands as well as texture and contextual measures for the concerned sub-image as the independent variables. Three types of texture were used: urban pixel density, homogeneity and entropy. The first one is simply defined as the proportion of all pixels within a sub-image classified as urban land-cover; the homogeneity is derived from a joint probability distribution ( $P$ ) matrix by measuring the sum of the squares of the matrix entries; and the entropy is also applied to a  $P$  matrix, yielding higher values when matrix entries are equal, and lower values when they are unequal. Only a single contextual measure was used: distance from the city centre. The results obtained indicated that the best fit was achieved when using TM bands 5 and 7 and the homogeneity measure only, since it is a normalized measure, and hence, its size is independent of the absolute number of urban pixels in a sub-image.

Nighttime satellite imagery has also been used for estimating the population density in the USA [3] and in China [4]. In the first work, the authors derived the nighttime stable-lights imagery from the visible near infrared band of 231 orbits of the Defense Meteorological Satellite Program Operational Linescan System (DMSP-OLS). Analyses of the saturated areas of the images indicated strong correlations between the areas of the saturated clusters and the populations those area cover.

The radiance-calibrated DMSP-OLS data was also used to evaluate their potential for population estimates in China at the provincial, county, and city levels [4]. The light clusters were classified into six categories of light intensity, and their areal extents were extracted from the image. Mean pixel values of light clusters corresponding to the settlements were also determined. The author conceived a light volume measure to indirectly assess the three-dimensional population capacity of a settlement. The author dealt with linear regression models, regarding the Chinese population and population densities as the dependent variable, and the light area, light volume, and pixel mean as independent variables. It was found that the DMSP-OLS nighttime data produced reasonably accurate estimates of urban population using either the light area or light volume as input. The total sums of the estimates for both urban and total population closely approximated the true values given by the Chinese statistics at all three spatial levels.

A pixel-based alternative approach for modeling population estimation was as well proposed [5]. In this method, pixels of a LS-TM image were classified as residential or non-residential using standard techniques. Initial reference populations were assigned by uniformly distributing the population of each zone across its residential pixels. An expectation-maximization (EM) algorithm was used to iteratively regress pixel population on spectral indicators and re-estimate pixel populations. This model produced population estimates of comparable accuracy to those resulting from a much more complex zone-based modeling procedure, and had its predictive validity evaluated by applying the fitted regression equation to a second image.

Recent works started to introduce high-resolution estimates for assessing the population spatial distribution. In this sense, the correlation between census population density and Ikonos image texture has been investigated [6]. The spatial unit adopted for the analysis was census blocks with homogeneous land use. Ikonos image texture was described using three methods: the gray-level co-occurrence matrix (GLCM), semi-variance, and spatial metrics. The authors suggest that the correlation is not strong enough to predict or forecast residential population, but the image texture does provide a base to refine census-reported population distribution using remote sensing.

This paper presents a new method for identifying homogeneous residential areas by means of an object-oriented multilevel segmentation and fuzzy classification of Quickbird images, developed in the knowledge-based platform e-Cognition. The identified homogeneous units are associated with certain ranges of population density, and based upon their areal extensions, it is possible to infer the amount of population living in each considered residential unit. Details regarding particularities of urban targets found in the urban scene and the methods employed

for this knowledge-based classification approach are given in the following sections.

## 2. THE STUDY AREA

The municipality of São José dos Campos is located in the southeastern State of São Paulo, within coordinates 46°06'W, 23°18'S, 45°40'W and 22°49'S. Its total surface amounts to 1,099.60 km<sup>2</sup> and it comprises a population of about 539,313 inhabitants, out of which 532,717 (98.78 %) inhabitants live in urban areas. It is situated along the road that connects São Paulo to Rio, and it is famous for being the country's aerospace pole. The city residential areas present very diverse spatial patterns, which correspond indeed to socio-economic standards of their dwellers. It is observable that the higher the standard of a given neighbourhood, the greater its share of urban greening, and hence, the lower its density of occupation. In the present work, only part of the southern sector of the city will be investigated. This selected study area, although of reduced size in comparison to the whole urban area of São José dos Campos, contains a great diversity of residential areas categories, ranging from very high, high, medium to medium-low and low standards. The area also contains three squatter settlements, popularly known as 'favelas'.

## 3. METHODOLOGY

### 3.1 Building the Database

Five bands of a Quickbird image were used in the database of this research: a panchromatic one, with 0.67 m of resolution, and four multispectral bands (blue, green, red, and near infrared) with 2.40 m of resolution. Their level of geometric correction is *ORStandard2A* and the acquisition date is May 17<sup>th</sup>, 2004. They present an off-nadir angle of 7.0° and radiometric resolution of 16 bits, although the pixels *DN* are actually distributed within a range of 11 bits.

Besides the image bands, one GIS layer related to the study area streets network (in shape file format) was imported to the database. This layer, issued by the planning department of the local government [7], was used for the initial segmentation of streets and blocks, which effectively guided the segmentation of further levels of the database.

For the classification of urban land cover of the study area, i.e. roof materials, paving materials, shadow, trees and grass areas, the database was initially structured in three levels of segmentation [8], in the following sequence: (i) Level 3 – segmented according to the GIS layer of the streets network, with scale factor of 10,000; (ii) Level 2 – segmented after Level 3, with scale factor of 50; and (iii) Level 1 – segmented after Level 2, with a very fine scale factor of 20. Such scale factor is an internal parameter of e-Cognition [9], and the user sets it heuristically in each segmentation level.

The classification of Level 3 corresponds to blocks and streets; Level 2, on its turn, was roughly classified in vegetation, shadow, and built-up areas; and finally, the classification of Level 1 refers to the urban land cover of the study area (Figure 1), including: trees, grass, shadow, asphalt, light and dark bare soil, light and dark ceramic roofs, dark concrete/asbestos roofs, medium tone concrete/asbestos roofs, metallic roofs, and swimming-pools.

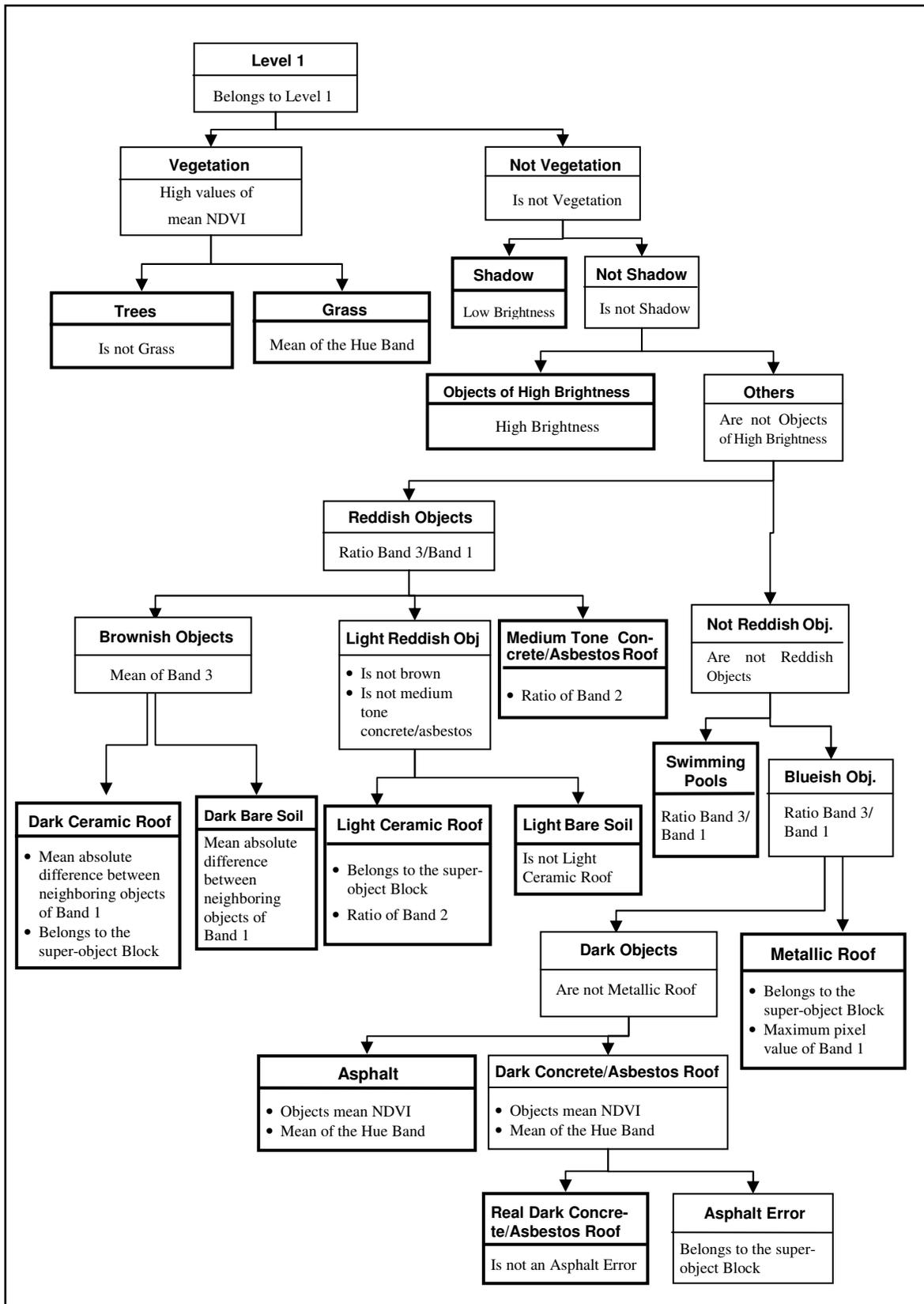


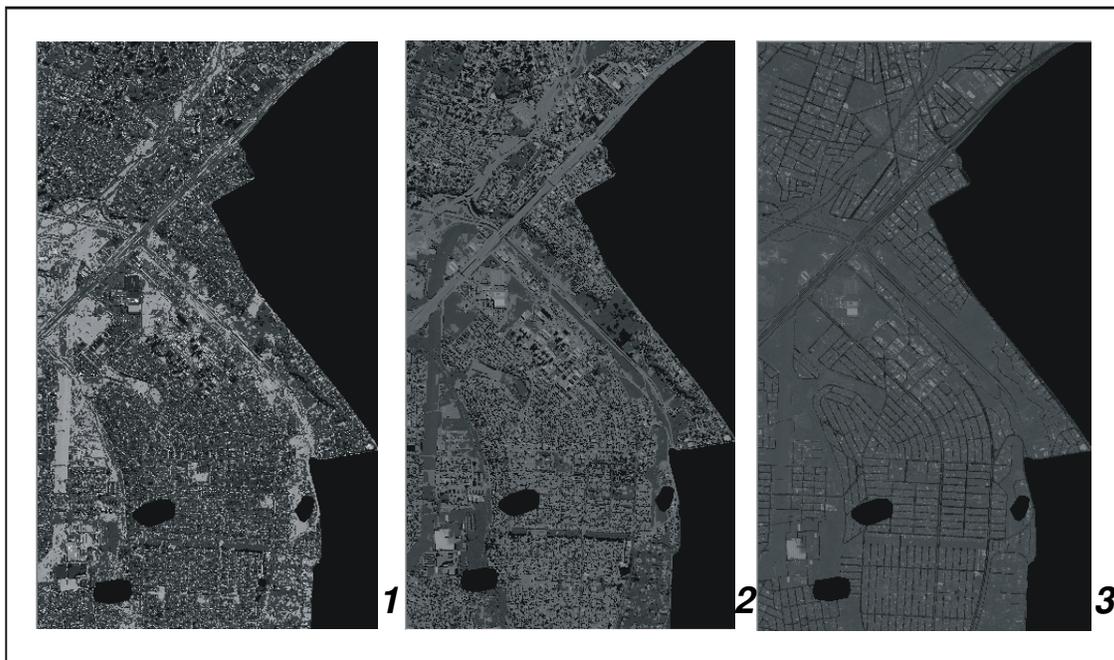
Figure 1. Hierarchical network of Level 1 and respective attributes for each class. Source: PINHO (2005).

Only one geometrical attribute was necessary to differentiate streets from blocks at Level 3: the rectangular fit. At Level 2, shadow was defined as a function of the objects brightness; vegetation was identified by means of the objects NDVI, and built-up areas were classified by mutual exclusion of these two former classes.

Each of the attributes of Level 1 presented a fuzzy membership curve, built upon samples collected on the image. Some of these curves were continuous, and others crisp. It is worthy reminding that the objects of Level 3 were decisive for the classification of Level 1. The final classification of Levels 1, 2 and 3 are presented in Figure 2.

- (iii) high standard houses of type I (big plot areas and high share of green areas);
- (iv) high standard houses of type II (very large plot areas and also high share of green areas);
- (v) medium standard houses of type I (ceramic roofs);
- (vi) medium standard houses of type II (asbestos roofs);
- (vii) low standard houses of type I (ceramic roofs);
- (viii) low standard houses of type II (asbestos roofs);
- (ix) squatter settlements ('favelas').

The final classification of these homogeneous units, undertaken at Level 4, took several attributes into consideration (Table 1), mainly found at Level 1.



**Figure 2. Classification of Levels 1 (urban land cover), 2 (vegetation, shadow and built-up areas) and 3 (streets and blocks). Black holes inside the classified images correspond to clouds or clouds shadow.**

### 3.2 Defining the Homogeneous Residential Units and their Respective Attributes

In order to classify the residential areas of the city into homogeneous units, the blocks had to be firstly classified into residential and non-residential areas. This was done manually at Level 5, since the local Planning Department of the city did not dispose of these data in digital format. Moreover, some residential areas occupied only part of the blocks, so there has been a vector edition, so as to subdivide blocks into residential and non-residential areas.

The next step was the definition of the homogeneous residential units. Nine classes of homogeneous residential areas were taken into account within the study area:

- (i) high-rise buildings of type I (single buildings);
- (ii) high-rise buildings of type II (sets of buildings);

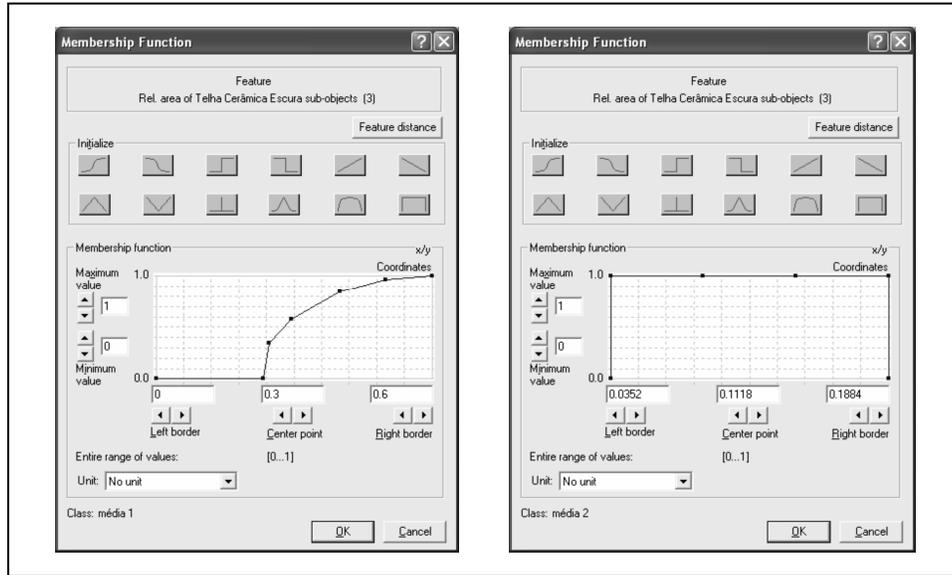
A specific fuzzy membership curve was created for each of the above attributes, considering the behavior of these attributes regarding the concerned class. Again, some curves were fuzzy, and others crisp. Figure 3 presents some illustrative examples of such curves in the case of the relative area of Dark Ceramic Roof for Medium Standard of types I and II.

### 3.3 Creation of a Customized Indicator of Streets Greening Density

The homogeneous residential units could be mostly differentiated by the attributes characterizing the share of certain types of roof material or the share of vegetated areas. A problem regarding the discrimination of medium and low standard residential areas became evident: the share of roof material and green spaces alone could not cope with the separation among these classes.

**Table 1. Attributes adopted to discriminate the homogeneous residential units at Level 4.**

<b>Classes</b>	<b>Attributes for the classes discrimination</b>
<b>High-Rise I</b>	Existence of Objects of High Brightness Relative area of Objects of High Brightness Relative area of Dark Concrete/Asbestos Roof Relative area of Shadow Relative area of Trees
<b>High-Rise II</b>	Existence of Objects of High Brightness Relative area of Objects of High Brightness Relative area of Dark Concrete/Asbestos Roof Relative area of Shadow Relative area of Trees
<b>High Standard I</b>	Total Area Relative Area of Trees Relative area of Dark Ceramic Roof
<b>High Standard II</b>	Total Area Relative Area of Trees Relative area of Dark Ceramic Roof
<b>Medium Standard I</b>	Relative area of Dark Ceramic Roof Relative area of Dark Concrete/Asbestos Roof Customized Indicator of Streets Greening Density
<b>Medium Standard II</b>	Relative area of Dark Ceramic Roof Relative area of Dark Concrete/Asbestos Roof Customized Indicator of Streets Greening Density
<b>Low Standard I</b>	Relative area of Dark Ceramic Roof Relative area of Dark Concrete/Asbestos Roof Customized Indicator of Streets Greening Density
<b>Low Standard II</b>	Relative area of Dark Ceramic Roof Relative area of Dark Concrete/Asbestos Roof Customized Indicator of Streets Greening Density
<b>Squatter Settlements (ˆFavelasˆ)</b>	Not existence of Swimming-Pools Relative area of Objects of High Brightness Relative area of Metallic Roof Relative area of Dark Concrete/Asbestos Roof Relative area of Shadow Relative area of Grass



**Figure 3. Fuzzy membership curves regarding the attribute “Relative area of Dark Ceramic Roof” (*Telha Cerâmica Escura*) for Medium Standard Residential Areas of Type I (left) and II (right). The curve to the left has a fuzzy configuration, whereas the one to the right presents crisp limits.**

One peculiar characteristic of medium standard and low standard residential neighbourhoods is the obvious difference as to their density of streets greening (streets trees density). Usually, low standard neighbourhoods arise irregularly. They are further regularized, but the streets greening initiatives by the local government are not as effective as those promoted in the case of formal medium standard neighbourhoods.

For assessing the streets greening density, a first solution was to estimate the total trees area inside a 15-meter buffer surrounding each residential block. A second possibility was to calculate the share of the intersection perimeter between the class “Streets” and the class “Residential Block” in relation to the total block perimeter. Both solutions were combined into one and they were implemented outside the e-Cognition platform, in the TerraView opensource software, developed by the Division for Images Processing of the Brazilian National Institute for Space Research (DPI-INPE)<sup>1</sup>. The reason for this external implementation lies on the fact that there is no buffer operation inside e-Cognition. Furthermore, the intersection perimeter can only be calculated for classes belonging to the same level inside e-Cognition. Regrettably, due to the huge number of trees polygons (over 102,000), a merging operation between the classes “Trees” and “Residential Block” was not rendered possible due to memory capacity constraints of such platform.

A threshold was established for the “Streets Greening Density Indicator”, designed to differentiate medium from low standard residential areas (Equation 1). This indicator was further

imported to the e-Cognition database at Level 6, indicating blocks above and below the threshold, as follows:

$$SGD = TA \leq 100 \text{ m}^2 \text{ AND } IP/TBP \leq 0.06 \quad (1)$$

where TA refers to the trees area inside the 15-m buffer, IP relates to the intersection perimeter between the polygons belonging to the class “Trees” (lying over the class “Streets”) and to “Residential Block”, and TBP corresponds to the total residential block perimeter.

### 3.4 Final Hierarchical Framework for the Classification of Homogeneous Residential Units

The process of extracting homogeneous residential units was organized upon basis of a double-stage knowledge-based classification process, well summarised in Figure 4.

Starting from a Quickbird image and a GIS layer containing the city streets network, assigned to Level 3, a classification of vegetation, shadow and built-up areas was carried out and assigned to Level 2. Observing these two levels, a classification of urban land cover was finally made at Level 1.

On a second stage, residential and non-residential blocks were manually classified at Level 5. A customised indicator of streets greening density was limiarised in a binary form and spatialised at Level 6. Simultaneously considering Levels 1, 5, and 6, a classification of homogeneous residential units was automatically accomplished at Level 4.

<sup>1</sup> <http://www.dpi.inpe.br/terraview.html>

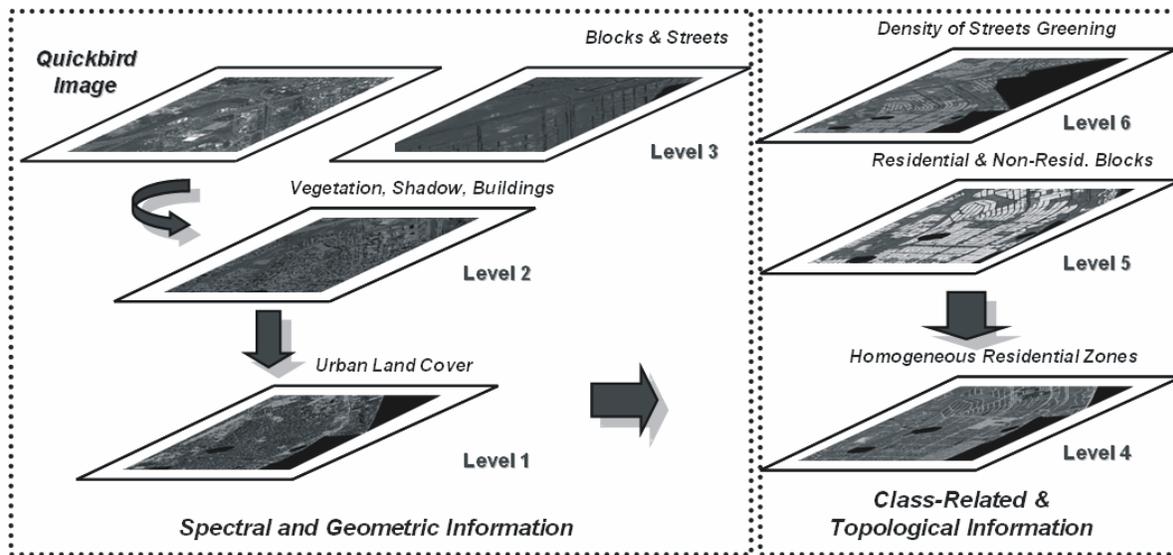


Figure 4. Schematic flowchart of the knowledge-based classification process of homogeneous residential units.

In this way, Levels 1 to 3 concern the first stage of the classification process, taking into account mainly spectral and geometrical information. Levels 4 to 6, on the other hand, regard a second stage of such classification process, mostly considering class-related and topological information.

#### 4. RESULTS AND DISCUSSION

The final classification of Levels 4, 5, and 6 are presented in Figure 5. It is noticeable that only one polygon remained non-classified at Level 4 and it in fact concerns a high standard residential zone under construction, which presents high percentages of Bare Soil and low percentages of Ceramic Roof.

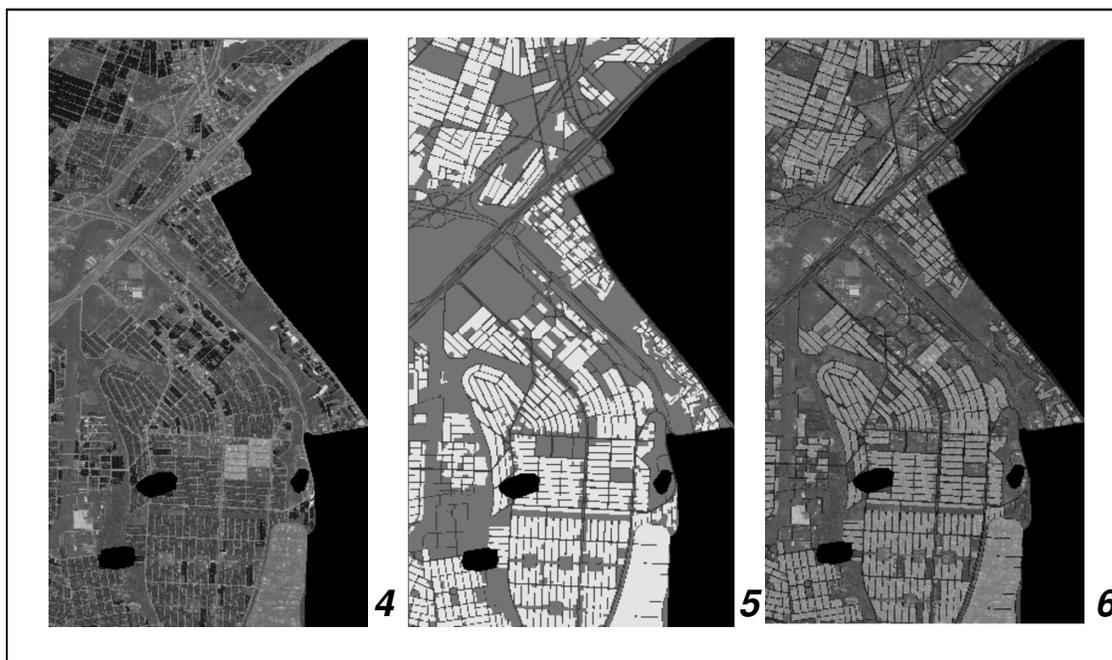


Figure 5. Final classifications of Level 4 (homogeneous residential units), Level 5 (residential and non-residential blocks), and Level 6 (indicator of streets greening density in binary form, showing blocks above and below the threshold set by Equation 1).

**Table 2. Validation: Contingency table for the polygons-based Kappa statistics.**

Ground Truth		HR1	HR2	HS1	HS2	MS1	MS2	LS1	LS2	FAV	NC	Total
Classification	HR1	66										66
	HR2		10									10
	HS1			1								1
	HS2				1							1
	MS1					376		16	1			393
	MS2					1	9		1			11
	LS1					48		65				113
	LS2								6			6
	FAV									3		3
	NC											1
<b>Total</b>		<b>66</b>	<b>10</b>	<b>1</b>	<b>1</b>	<b>425</b>	<b>9</b>	<b>81</b>	<b>8</b>	<b>3</b>	<b>1</b>	<b>605</b>

The High-Rise Building Areas were well identified, and the attributes for these two types are exactly the same, only differing in the fuzzy curves and fuzzy membership values. The High Standard Unifamiliar (Single Dwelling Units) Zones were also well discriminated upon basis of the neighbourhoods area, relative area of Trees and of Dark Ceramic Roof.

The Medium Standard Residential Area of Type I is the most common residential unit in the city, with a maximum of 25% of Asbestos Tiles, and a minimum of 30% of French Tiles. On the other hand, Standard Residential Area of Type II concerns a very particular neighbourhood in the city, and presents from 40% to 50% of Asbestos Tiles, and up to 18% of French Tiles.

Although the 'Favelas' have the most complex set of attributes, they were well identified. The greatest confusion remained in the discrimination between Medium Standard from Low Standard Residential Areas, of both Types I and II.

## 5. CONCLUSIONS AND DIRECTIONS FOR FUTURE WORK

The final classification of homogeneous residential units at Level 4 was subject to validation by means of a ground truth level (Level 7) built upon field observation. The contingency table is presented below (Table 2), and the value of the polygons-based Kappa statistics reached 78%. According to [10], this rate can be regarded as very good, since the classification quality is set according to the ranges of the Kappa index as: [ $<0.00$ ]: very bad; [ $0.00 - < 0.20$ ]: bad; [ $0.00 - 0.40$ ]: acceptable; [ $0.40 - < 0.60$ ]: good; [ $0.60 - < 0.80$ ]: very good; [ $0.80 - 1.00$ ]: excellent.

The authors acknowledge this is a work in progress, and will continue on tackling many open questions and mismatches in the classification process. As directions for future work, the authors plan to investigate texture measures related to specific classes of homogeneous residential units, to explore further customized (arithmetic and/or relational) attributes, to insert census and cadastral data (plots and housing units sizes) as well as to include 3-D models derived from LIDAR data, in order to precisely assess the number of stores in high-rise buildings, and hence, support more consistent methods for population estimates.

## 6. ACKNOWLEDGMENTS

The authors would like to express their gratitude to the Brazilian National Council for Scientific Research (CNPQ) through grant number 05/06-491084.

## 7. REFERENCES

- [1] Langford, M. and Unwin, D. J. Generating and mapping population density surfaces with a geographical information system. *The Cartographic Journal*, 31 (1994), 21-26.
- [2] Webster, C. J. Population and dwelling unit estimates from space. *TWPR*, 18, 2 (1996), 155-176.
- [3] Sutton, P., Roberts, D., Elvidge, C., and Meij, H. A comparison of nighttime satellite imagery and population density for the continental United States. *Photogrammetric Engineering & Remote Sensing*, 63, 11 (1997), 1303-1313.
- [4] Lo, C. P. Modeling the population of China using DMSP Operational Linescan System nighttime data. *Photogrammetric Engineering & Remote Sensing*, 67, 9 (2001), 1037-1047.
- [5] Harvey, J. T. Population estimation models based on individual TM pixels. *Photogrammetric Engineering & Remote Sensing*, 68, 11 (2002), 1181-1192.
- [6] Liu, X., Clarke, K., and Herold, M. Population density and image texture: a comparison study. *Photogrammetric Engineering & Remote Sensing*, 72, 2 (2006), 187-196.
- [7] PMSJC - Prefeitura Municipal de São José dos Campos. 2003. *Cidade Viva*. [CD-ROM], São José dos Campos: PMSJC, 2003.
- [8] Pinho, C. M. D. *Análise orientada a objetos de imagens de alta resolução espacial*. São José dos Campos. 2005. Master's Thesis - National Institute for Space Research.
- [9] Definiens. 2003. *eCognition: User Guide 3*. 480 p. Available at: <http://www.definiens-imaging.com/down/ecognition>. Access on: 15. Oct., 2005.
- [10] Landis, J. R. and Koch, G. G. The measurement of observer agreement for categorical data. *Biometrics*, 33, 1 (1977), 159-174.