

# A containment-first search algorithm for higher-order analysis of urban topology

J.-P. de Almeida<sup>a, b, 1</sup>, J. G. Morley<sup>c</sup>, I. J. Dowman<sup>d</sup>

<sup>a</sup> Geomatic Engineering Unit, Dept of Mathematics, Faculty of Science and Technology, University of Coimbra (FCTUC), Largo D. Dinis, Apartado 3008, 3001-454 COIMBRA, Portugal  
Tel. +351 239 791 150 Fax +351 239 832 568

[zepaulo@mat.uc.pt](mailto:zepaulo@mat.uc.pt)  
<http://www.mat.uc.pt>

<sup>b</sup> Institute for Systems Engineering and Computers at Coimbra (INESCC),  
Rua Antero de Quental, 199, 3000-141 COIMBRA, Portugal  
Tel. +351 239 851 040/9 Fax +351 239 824 692

<http://kreation.dec.uc.pt/projects>

<sup>c</sup> Centre for Geospatial Science, School of Geography, University of Nottingham (UN),  
Sir Clive Granger Building, University Park, NOTTINGHAM NG7 2RD, UK  
Tel. +44 (0)11 5846 8411 Fax +44 (0)11 5951 5249

[jeremy.morley@nottingham.ac.uk](mailto:jeremy.morley@nottingham.ac.uk)  
<http://www.nottingham.ac.uk/geography/CGS/>

<sup>d</sup> Dept of Civil, Environmental & Geomatic Engineering, University College London (UCL),  
Gower Street, LONDON WC1E 6BT, UK  
Tel. +44 (0)20 7679 2740 Fax +44 (0)20 7380 0453

[idowman@cege.ucl.ac.uk](mailto:idowman@cege.ucl.ac.uk)  
<http://www.cege.ucl.ac.uk>

**Keywords:** topology, Graph Theory, urban; scene analysis, GIS.

---

## 1. Introduction

### 1.1 Motivation

Research has revealed the importance of the concepts from the mathematical areas of both topology and graph theory for interpreting the spatial arrangement of spatial entities. Graph theory in particular has been used in different applications of a wide range of fields for that purpose, however not many graph-theoretic approaches to analyse entities within the urban environment are available in the literature. Some examples should be mentioned though such as, Bafna (2003), Barr and Barnsley (2004), Bunn *et al.* (2000), Krüger (1999), Nardinochi *et al.* (2003), and Steel *et al.* (2003).

Very little work has been devoted in particular to the interpretation of initially unstructured geospatial datasets. In most of the applications developed up-to-date for the interpretation and analysis of spatial phenomena within the urban context, the starting point is to some extent a

---

<sup>1</sup> Corresponding author.

meaningful dataset in terms of the urban scene. Starting at a level further back, before meaningful data are obtained, the interpretation and analysis of spatial phenomena are more challenging tasks and require further investigation.

The aim of retrieving structured information from initial unstructured spatial data, translated into more meaningful homogeneous regions, can be achieved by identifying meaningful structures within the initial random collection of objects and by understanding their spatial arrangement (Anders *et al.*, 1999). It is believed that the task of understanding topological relationships between objects can be accomplished by both applying graph theory and carrying out graph analysis (de Almeida *et al.*, 2007).

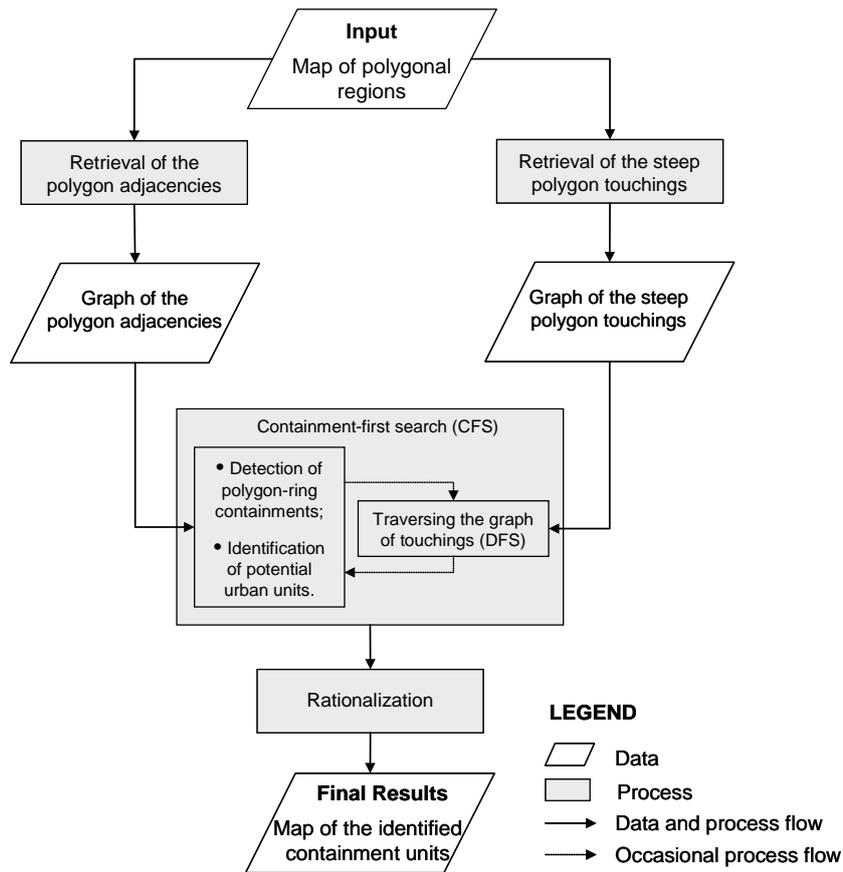
## 1.2 Background

Starting from initially unstructured geospatial datasets of urban areas (thus, no prior knowledge of the spatial entities is assumed), de Almeida *et al.* (2005, 2007) showed how a graph-theoretic approach could be applied towards the analysis of the urban scene spatial topology.

Urban LiDAR data was used as an example scenario. Topology was initially brought in to the original data by generating a triangulated irregular network (TIN - the maximal planar description of the given point set's internal structure, Kirkpatrick and Radke, 1985). A binary classification of the TIN facets based upon their gradient – whose thresholding depends on the resolution of the initial data – was employed (“flat” and “steep” facets). Eventually, the TIN facets were “aggregated” according to the classification above that led to a map of polygonal gradient regions (“flat” and “steep” polygons). The authors pointed out how the steep polygonal regions in particular were expected to enclose urban features. These steps constitute the preliminary preparation process of raw data.

A network of connectivity throughout the map of flat & steep polygonal regions was then built up by applying graph theory, which resulted in a graph of adjacencies: each region in the graph is represented by a node; graph edges link up nodes corresponding to adjacent polygons. The adjacency graph was processed either through the depth-first (DFS) or the breadth-first search (BFS) algorithms. Given the different ways each algorithm operates in traversing a graph, it was noted how BFS results are more meaningful in terms of the urban scene: the BFS tree branches are connected components of the original graph, and represent the shortest path between the root and their leaf (Sedgewick, 2002); it seems that they can be related to potential urban features. Thus, the implementation of the graph analysis procedure was based upon BFS. It traverses the graph looking for sequential relationships of containment amongst the sequences of adjacency: containment-first search (CFS). In fact, where containment occurs within the Useful External Border (UEB) – basically, the outer flat enclosing polygon corresponding to the ground – there is a high likelihood of an urban feature being present (de Almeida *et al.*, 2005, 2007).

This paper describes in detail CFS algorithm which was developed for the purposes above. The diagram depicted in Figure 1 above illustrates where the algorithm sits within the whole methodology proposed for the analysis of urban spatial topology.



**Figure 1.** An overview of the methodology proposed for the analysis of urban spatial topology (de Almeida, 2007).

## 2. A containment-first search algorithm

### 2.1 Preliminaries

For the purposes of this work, when two polygons share at least one arc, the spatial relation is called *adjacency*; if the two polygons happen to meet at a node, the spatial relation is distinguished from the previous one and is called *touching* (de Almeida *et al.*, 2007; de Almeida, 2007).

As de Almeida *et al.* (2007) noted, CFS could not be developed simply based on BFS but had to be extended in order to be able to detect the spatial relation of containment in a broader sense. The spatial relation of *touching* between steep polygons should be taken into consideration. This improvement enabled the derivation of particular cases of containment not explicit in the graph of adjacencies, e.g. when a ring of steep polygons meeting at nodes contains a single flat polygon - the so-called polygon-ring containments.

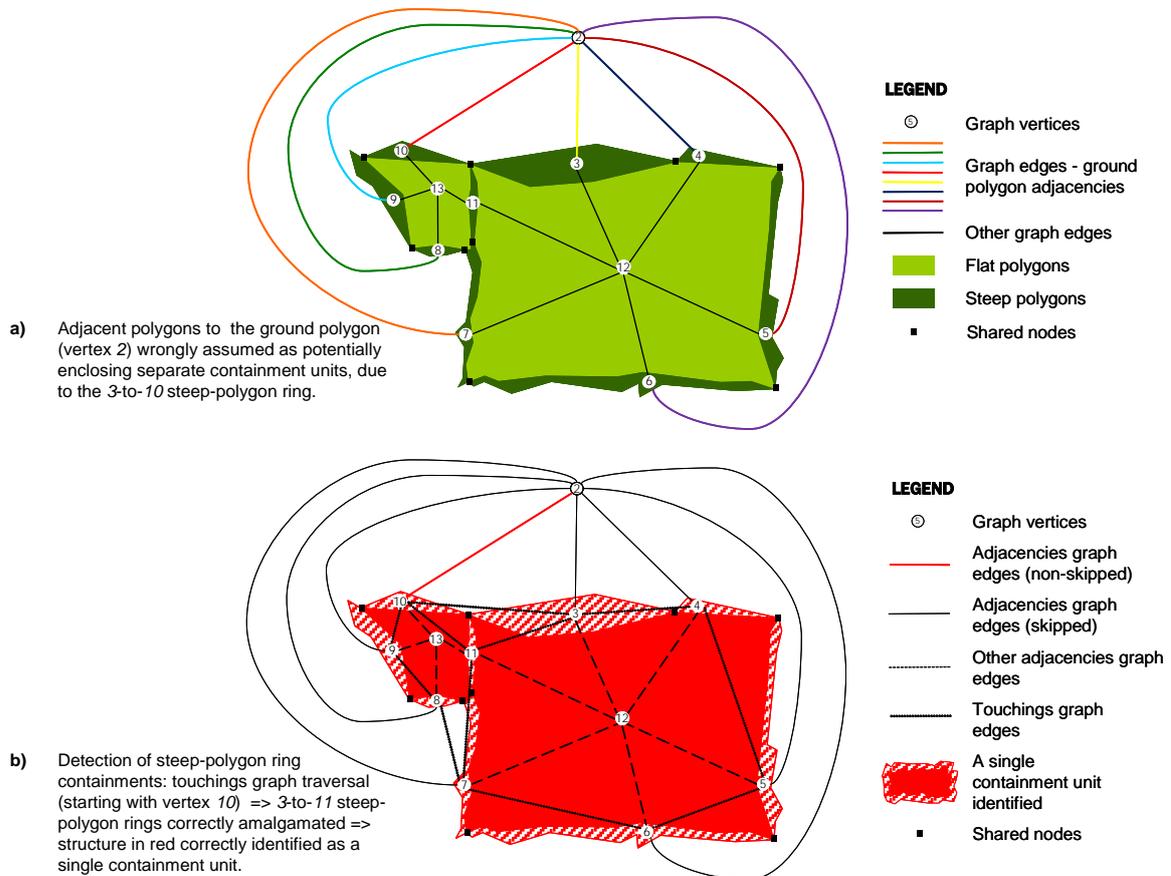
### 2.2 The analytical analysis method

The graph-theoretic approach was implemented based on the investigation of the topological relationships between objects in the context of the whole spatial scene. This was coded in C foreseeing the advantages and potentialities of pointer structures in C for graph analysis (Kelley and Pohl, 1990).

The analysis method can be interpreted as follows. Considering the UEB (recall that this is a flat polygon) as the starting point of the search process, the original graph of adjacencies is

traversed. When visiting the adjacent steep vertices of the root, the CFS algorithm takes the first vertex appearing in the root's linked list and, starting from this one, traverses the graph of steep-polygon touchings. Because the graph of touchings is a disconnected graph, the traversal process covers only the subgraph that the given steep vertex belongs to. While traversing this particular subgraph, the CFS algorithm tags all the steep vertices visited as belonging to the same connected unit. This process continues until the first level of adjacency of the graph of adjacencies is exhausted. When the CFS comes across a root's adjacent vertex already tagged as belonging to a particular containment unit, this is skipped and the corresponding polygon remains intact, belonging to the containment unit already identified.

To illustrate the concept implemented, let us take a simpler scene pictured in Figure 2. Let us suppose that steep polygons 3,...,11 (in dark green) are constituent parts of the rings of steep polygons enclosing flat polygons 12 and 13 (vd. Figure 2a); in other words, there is a sub-graph of the graph of steep-polygon touchings that consists of vertices 3 to 11.



**Figure 2.** The containment-first search process: **a)** before polygon-ring containments are detected; **b)** after polygon-ring containments are detected

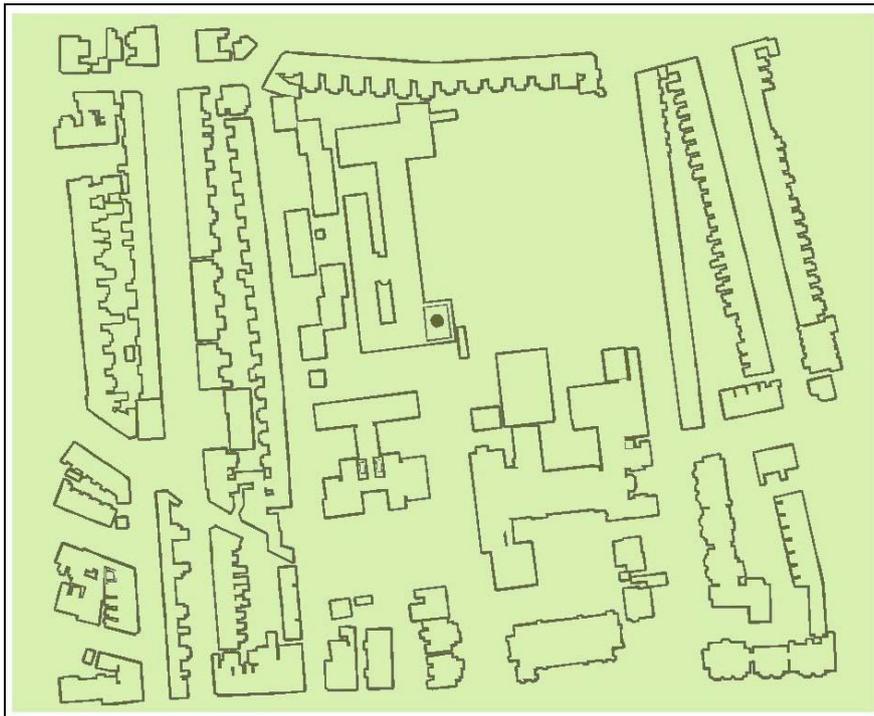
When vertex 2 is visited in the adjacencies graph, the algorithm takes the vertex at the top of 2's adjacency list, vertex 10, and the graph of steep-polygon touchings is traversed starting from 10; all the steep vertices belonging to the same sub-graph as that of 10 are tagged accordingly, indicating a potential containment unit. When vertex 10 is exhausted in the graph traversal, CFS moves on to visit vertex 9; this is now skipped since it was previously tagged as belonging to an containment unit already identified. And so on so forth until vertex 3 is visited, and the containment unit is complete. Visually, the translation of the facts above is accomplished by assigning the same colour to all steep (hashed pattern) and flat (solid colour) polygons within the same containment unit (vd. Figure 2b).

### 3. Proof of concept

#### 3.1 Generation of synthetic data

Before tests with real initial unstructured urban data are undertaken, this section describes an experiment carried out with synthetic idealised spatial data relating to urban objects. A map of binary classified gradient regions was created simulating a map of higher-level urban scene objects. This was derived from building polygons from OS Master Map data<sup>2</sup>.

As Figure 3 shows, steep polygons shape both buildings standing on their own and higher-level structures; the enclosed flat polygons simulate building roofs. For topological reasons, an outer polygon - distinct from the Universe Polygon - had to be considered so as to simulate the ground polygon.



**Figure 3.** Simulated map of gradient regions: binary classification of the polygons generated into “steep” polygons (dark colour) and “flat” polygons (light colour).

Polygon and associate arc attributes were accessed in order to retrieve gradient-region adjacencies (ESRI, 1995; Rigaux *et al.*, 2002; ESRI, 2005). The graph of adjacencies was generated. The main characteristic of this graph when dealing with clean data is the fact that all relationships of adjacency are also of containment. This is far more complex when dealing with real world data: it is not guaranteed beforehand that a spatial relation of adjacency is also of containment (de Almeida *et al.*, 2007).

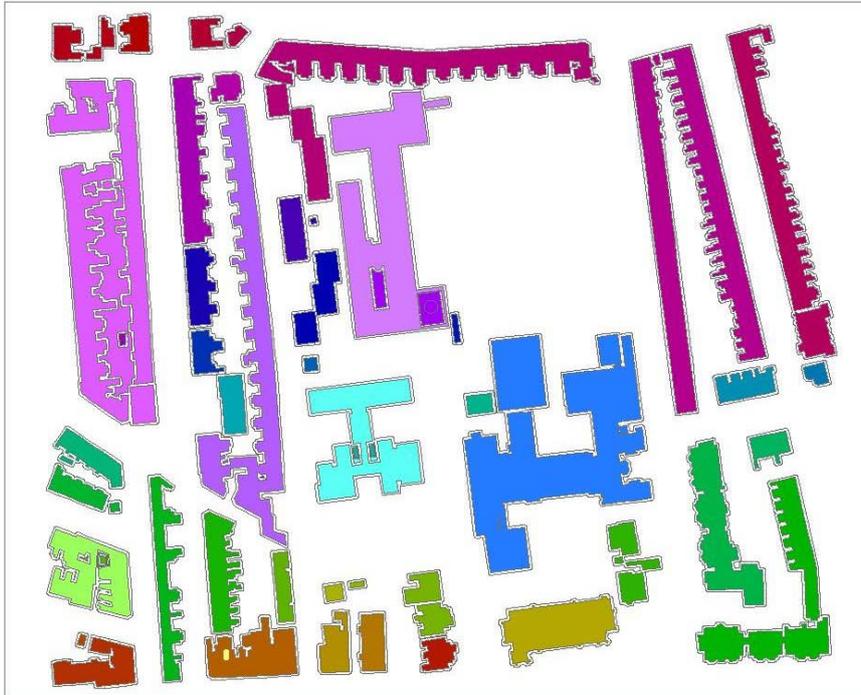
As typically happens with real world data, there are no steep polygons meeting at nodes in this case (*i.e.* steep-polygon rings, enclosing single containment units, are not split into different entities), and hence the touchings graph is a null graph.

---

<sup>2</sup> Made available by the Department of Geomatic Engineering of the University College London for academic purposes. Ordnance Survey ©Crown Copyright, all rights reserved.

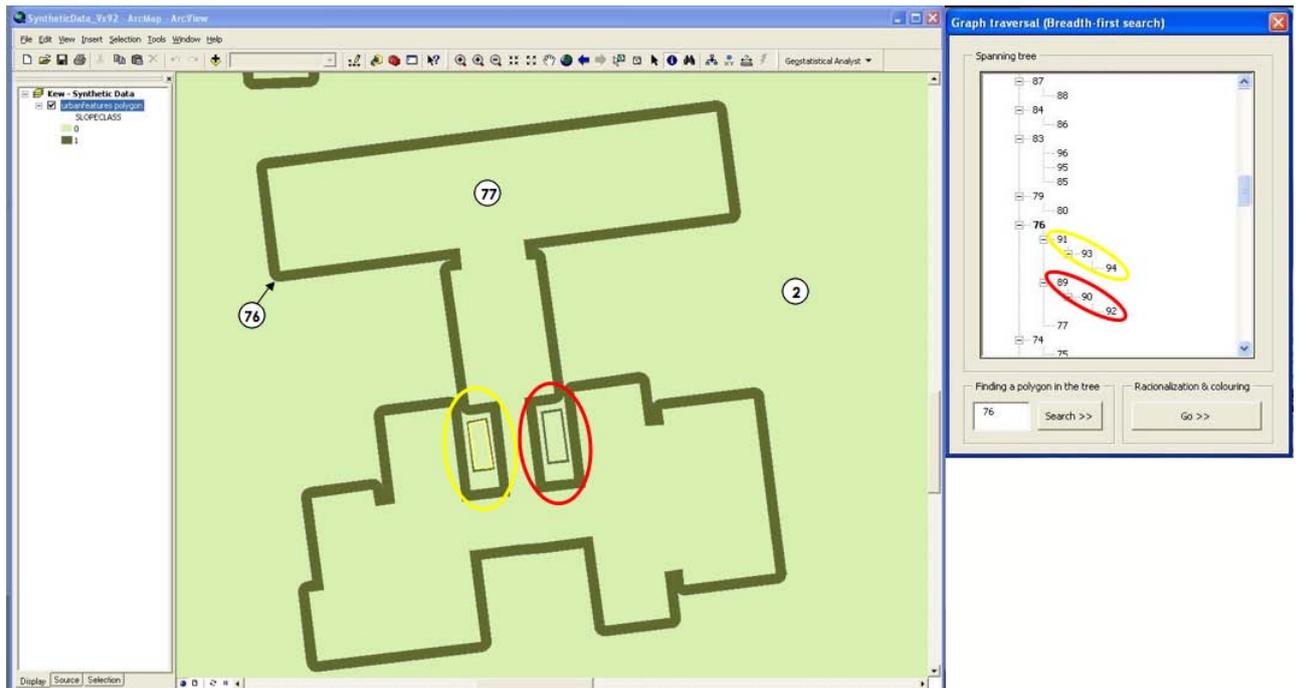
### 3.2 Spatial topology analysis

Figure 4 depicts the results of the spatial topology analysis for the simulated map of gradient regions. Polygon 2 (corresponding to the ground polygon, mapped in white), with 45 adjacent regions, was chosen as the UEB.



**Figure 4.** Spatial topology analysis and the different containment units identified.

It can be seen that the algorithm indeed detected individually all the urban features simulated as separate containment units (different colours represent each one). In fact, sequences of adjacencies/containments were correctly detected as so by the algorithm (*vd.* Figure 4): solid colours correspond to flat polygons; coloured hashed patterns correspond to steep polygons. Moreover, individual simulated spatial features, closely standing next to one another but not actually juxtaposed, were detected separately. This confirms that in theory the algorithm is even capable of detecting single buildings standing on their own.



**Figure 5.** Branches of the breadth-first search tree directly relate to spatial features. (Detail of Figure 3; the numbers on the map are polygon labels in insert)

#### 4. Summary and conclusions

Further to our prior work, this paper showed how the spatial relation of touching between steep polygons was taken into consideration in order to extend the CFS procedure to be able to detect polygon-ring containments. A flowchart of the algorithms implemented was provided, and an illustration of how CFS procedure works was also given.

Proof of concept was carried out. For the purpose, synthetic data was generated and a map of gradient regions, simulating urban scene objects, was created. The analysis of the spatial topology was undertaken, and conclusions were drawn in terms of the assertions made when designing the algorithms. The results obtained demonstrated that, in the absence of noise and error, the algorithms do indeed make the urban spatial topology more explicit. In particular, the results support the assumption that each BFS tree's branch does relate to a single containment unit within the initial map of gradient regions (*e.g.* tree branch starting at vertex 76 in Figure 5). Moreover, sequences of containment relationships do relate to higher-level urban scene objects.

The concepts drawn in the research described and the algorithm implemented should serve as the basis for automatic analysis of spatial datasets, such as: analysis of image data; analysis of settlement structures; automation of land use mapping for urban areas. Furthermore, the results obtained particularly with LiDAR data reveal that the methodology proposed has also promised as a tool that could be extended to be applied in the automatic classification of raw LiDAR data.

#### References

- de ALMEIDA J.-P., MORLEY J. G., DOWMAN I. J. (2005). A graph-based technique for higher order topological data structure visualisation. In *Proceedings of GIS/UK 2005*: 306-312.

- de ALMEIDA J.-P., MORLEY J. G., DOWMAN I. J. (2007). Graph theory in higher order topological analysis of urban scenes. In *Computers, Environment and Urban Systems* **31**(4): 426-440.
- de ALMEIDA J.-P. (2007). *A Graph-based Technique for Analysis and Visualisation of Higher Order Urban Topology*. PhD thesis (unpublished). Department of Geomatic Engineering, University College London, London (UK).
- ANDERS, K.-H.; SESTER, M.; FRITSCH, D. (1999). Analysis of settlement structures by graph-based clustering. *Semantische Modellierung, SMATI 99*: 41-49, Munich, (Germany).
- BAFNA, S. (2003). Space Syntax – A brief introduction to its logic and analytical techniques. *Environment and behaviour* **35**(1): 17-29. Sage Publications.
- BARR, S.L., BARNESLEY, M.J. (2004). Characterising the structural form of an urban system using built-form connectivity model concepts. *Spatial Modelling of the Terrestrial Environment* (KELLY, R.E.J., DRAKE, N.A., BARR, S.L., Eds.): 201-226. John Wiley, Chichester (England, UK).
- BUNN, A.G., URBAN, D.L., KEITT, T.H. (2000). Landscape connectivity: A conservation application of graph theory. *Journal of Environmental Management* **59**: 265-275. Academic Press (USA)
- ESRI (1995). *Understanding GIS, The ARC/INFO Method*, Lesson2. ESRI, Inc. Cambridge, England (UK).
- ESRI (2005). ArcGIS Topology [online]. Available from <http://www.esri.com/products/support/ArcGIS%20UK%20EXT/Topology.asp> [Accessed 10 June 2005].
- KELLEY A., POHL I. (1990). *A Book on C - Programming in C* (2<sup>nd</sup> Edn.). Benjamin Cummings, Redwood City, CA (USA).
- KIRKPATRICK, D.G., RADKE, J.D. (1985). *A framework for computational morphology. Computational Geometry* (TOUSSAINT, G.T., Ed.): 217-248. North-Holland, Amsterdam (The Netherlands).
- KRÜGER, MÁRIO (1999). On node and axial maps: distance measures and related topics. *European Conference on the Representation and Management of Urban Change*: 28-29 September 1999. Cambridge (England, UK).
- NARDINOCHI, C., FORLANI, G., ZINGARETTI, P. (2003). Classification and filtering of laser data. *ISPRS Workshop on 3D reconstruction from airborne laser scanner and InSAR data*. Dresden (Germany).
- RIGAUX J., SCHOLL M., VOISARD A. (2002). *Spatial Databases with Applications to GIS*, **8**. Morgan Kaufmann, San Francisco, CA (USA).
- SEDGEWICK. R. (2002). *Algorithms in C, Part 5* (3<sup>rd</sup> Edn.), **17, 18**. Addison-Wesley, Reading, MA (USA).
- STEEL, A.M., BARNESLEY, M.J., BARR, S.L. (2003). Inferring urban land use through analysis of the spatial composition of buildings identified in LiDAR and multispectral image data. *Remotely-Sensed Cities* (MESEV, T.V., Ed.). Taylor & Francis, London (England, UK).

## Acknowledgements

*Faculdade de Ciências e Tecnologia da Universidade de Coimbra* (Faculty of Science & Technology of the University of Coimbra, Portugal), and *Fundação para a Ciência e*

*Tecnologia* (The Science & Technology Foundation – Portuguese Ministry of Science, Technology and Higher Education - Scholarship SFRH/BD/9909/2002), which supported financially this research.

An acknowledgement is also due to Dr. João Coutinho-Rodrigues (Department of Civil Engineering, Faculty of Science and Technology, University of Coimbra, Portugal), J-P de Almeida's PhD co-supervisor at the University of Coimbra.

## **Biography**

**José-Paulo de Almeida** (*Grad Geomatic Engineering UC; MSc Civil Engineering - Specialisation Urban Engineering UC; PhD Geomatic Engineering UCL*) has been working at UC where he is currently Lecturer of Geomatic Engineering; he is also a junior researcher at INESCC. He's been working on: the interpretation of unstructured geospatial data in GIS environment using Graph Theory; spatial decision support systems; currently he's also interested on the semantic enrichment of 3D data towards the development of 3D city models.

**Jeremy Morley** (*MA Natural Sciences U. Cambridge; MSc Remote Sensing UCL & Imperial C*) is Deputy Director of the Centre for Geospatial Science (CGS) at the University of Nottingham. He was programme director of UCL's MSc in GIS (1998-2004) and of its BEng/MEng in Geoinformatics (2005-9). Over the last 15 years his research has focussed on the mapping of Mars in support of geological analysis; terrain mapping from LiDAR and InSAR; GIS interoperability and mashup WebGIS systems. He's been UCL's technical representative to the Open Geospatial Consortium since 2004.

**Ian Dowman** (*BSc Geography UCL; PhD Photogrammetry UCL; FRICS - Fellow of the Royal Institution of Chartered Surveyor*) has spent most of his career at UCL, where he's currently Professor of Photogrammetry and Remote Sensing. He's been involved in ISPRS for many years as a working group chair, commission president, and as Secretary General of ISPRS. He's also been chair of the sub group on Terrain Mapping of the CEOS Working Group on Calibration and Validation. He worked with aerial and close range photography, on the application of satellite data to mapping, and is currently involved in automatic feature extraction and geometric fusion of different types of data.