

Zero Cost Solutions of Geo-Informatics Acquisition, Collection and Production for Natural Disaster Risk Assessment

Z. C. Zou ^a, X. G. Lin ^b

^a *Department of Urban Planning, School of Architecture, Harbin Institute of Technology, Harbin, China*

^b *CSIRO Computational Informatics, Australian Capital Territory*

Email: Xunguo.Lin@csiro.au

Abstract: The global and local impacts of climate change and natural disasters cause great casualties and economic losses each year. Catastrophic events bring more damage to urban areas than rural places because of the high density of human population, property and the infrastructures. Geo-informatics is critical for all stages of disaster management including disaster prevention, preparedness, emergency decision making, disaster relief, rebuilding and recovery. It is critical for risk assessment of natural disasters.

Since all hazard risk components are location-specific and vary spatially, accurate risk assessment relies on geo-informatics. Geo-informatics as the basis of decision-making information has proved to be critical and essential to natural, technological and manmade disaster risk assessments.

Commercial sources of geo-informatics are usually expensive, especially for those developing countries and regions where living standards are low but natural disasters occur more frequently and cause large losses. The expense of geo-informatics acquisition, collection and production varies vastly with different update frequency, imagery quality, and the total amount of information carried. In contrast with commercial sources, free sources of geo-data and geo-software have limited comparable features. However, they have the flexibility to add new features by further development.

Based on a current study of urban natural disaster risk assessment and vulnerability mapping which aims at improving community safety and resilience to natural hazards, this paper shares our experience on zero cost solutions of geo-informatics acquisition, collection, and semi-automatic production procedures by using free internet resources of Google Maps, Google Earth, and free and/or open source software including QGIS (Quantum GIS), GRASS, SAGA, Monteverdi, Sextante GIS and Orfeo Toolbox. In order to enrich attributes to meet the requirements of risk assessment, methods of combining the acquired geo-informatics with other formats of geo-data (*e.g.*, Digital Elevation Model, raster layers, vector layers, and Microsoft Excel files) are introduced.

The purpose of this paper is to share our experience on zero-cost geo-informatics acquisition, collection and production solutions with urban planning managers, emergency managers, risk researchers and people with similar concerns. The geo-informatics developed from this research forms a basis for disaster vulnerability measurement which is an important part for urban disaster risk assessment.

Keywords: *Geo-informatics, GIS data acquisition, GIS data production, natural disasters*

1. INTRODUCTION

One of the priority purposes of the Hyogo Framework for Action 2005-2015 for disaster risk reduction is the use of knowledge, innovation and education to build a culture of safety and resilience at all levels of government (Zschocke and León, 2010). According to the United Nations International Strategy for Disaster Reduction (UNISDR) disaster data and statistics report, in the past three decades (1980 – 2011) 3455 floods, 2689 storms, 470 droughts, and 395 extreme temperature events occurred around the world. In the last 12 years (2000 - 2011), natural disasters have caused losses of about 1.3 trillion USD, affected 2.7 billion people, and killed 1.1 million people (UNISDR, 2012).

Geo-informatics is critical for disaster risk reduction activities and emergency decision making. This paper shares our practical experience on geo-informatics collection, acquisition and production based on a research project which aims at urban natural disasters risk assessment and vulnerability mapping. Geo-informatics acquired by the methods introduced in this paper is fundamental for the vulnerability mapping. For instance, bushfire is one of the main natural disasters in Canberra (called the Bush Capital). Bushfire risk to urban systems is closely related to geophysical vulnerability which is determined by many factors including local topography, vegetation distribution, urban layout, individual infrastructure segment and facility location, and fire behavior. Risk assessment and vulnerability evaluation for a whole study area and its sub-areas are based on the respective geo-information about all entities of value on the ground.

The geo-information used in our research is derived from methods introduced in this paper. For example, urban topography models are based on 90m Digital Elevation Model (DEM) data, the majority vegetation spatial models are from Australian Bureau of Statistics data, some infrastructures spatial models (e.g., road network) are based on OpenStreetMap data and facility spatial modelling (e.g., most houses and buildings) are all from our generated data.

The rest of this paper is organized as follows. In Section 2, internet-accessible free geo-data sources and open source geo-software are listed. In Section 3, a collection of free maps from OpenStreetMap (OSM) are introduced. In Section 4, acquisition of Google satellite images is illustrated. In Section 5, a pixel-based Support Vector Machine (SVM) supervised classification procedure for satellite imagery feature detection and extraction is demonstrated. Finally, in Section 6, a method to enrich attributes from raster layers, vector layers, and MS Excel files is introduced. The paper concludes in Section 7.

2. INTERNET-ACCESSIBLE FREE GEO-DATA AND OPEN SOURCE GEO-SOFTWARE

2.1. Internet-Accessible Free Geo-data

The development of IT and internet benefits geo-data collection, dissemination and acquisition. Official governments' websites from many countries publish their national geo-information to the general public. Information completeness and data precision depend on location. Also there are websites that publish global geo-data for public use. An incomplete list of internet-accessible global scale free geo-data is shown in Table 1 with various copyright conditions.

2.2. Free and Open Source Geo-software

Free and open source geo-software refers to software or packages licensed under free software licenses and/or open source licenses. Most source code of open source geo-software and packages are accessible and modifiable by users. Some of the free and open source geo-software and packages are listed in Table 2. Google Earth, Quantum GIS, GRASS, SAGA, Sextante GIS, and Orfeo Toolbox (OTB)¹ were used in our research.

3. FREE MAPS COLLECTION FROM OPENSTREETMAP

OSM is a project to build up a free geographic database of the whole world. It aims to build up a record of every single geographic object on the planet eventually and this began with mapping streets. Beyond street level, it includes administrative boundaries, bus routes, footpaths, buildings, waterways, pipelines, details of land use, beaches, post-boxes and even individual trees, depending on regions (Bennett, 2010). The OSM's aim is to create a set of world map data by crowdsourcing² which is free to use, editable, and licensed under new copyright schemes (Haklay and Weber, 2008). The worldwide available location information was

¹ OTB is distributed as an open-source library of image processing algorithms. Its official home page: <http://www.orfeo-toolbox.org/otb>.

² Crowdsourcing is the practice of obtaining needed services, ideas or content by soliciting contributions from a large group of people, especially from the online community rather than from traditional employees or suppliers.

gathered by driving, cycling, or walking along streets and paths, and around areas, based on affordable GPS receivers. Geo-information was disseminated by mobile phones, pads, computers and the internet. To ensure the accuracy of location information provided by OSM, massive individual contributions both in processing technologies and in data structures have enhanced the OSM data quality (Coast, 2011; Girres and Touya, 2010; Lin, 2011; Ying *et al.*, 2010). Apart from the individual contributions, the development of OSM has also been supported by commercial GIS companies including ESIR (ITN, 2012; PR Newswire, 2012).

Table 1: List of internet-accessible global range free geo-data.

Type	Name and Hyper Link	Last Update
Vector Data	Global Self-consistent, Hierarchical, High-resolution Shoreline Database (GSHHS) http://www.ngdc.noaa.gov/mgg/shorelines/data/gshhg/latest/	Feb 22, 2013
DEM	Global Land One-kilometer Base Elevation (GLOBE) http://www.ngdc.noaa.gov/	Oct 14, 2011
DEM	Global 30 Arc Second Elevation Dataset (GTOPO30) http://webmap.ornl.gov/wcsdown/dataset.jsp?ds_id=10003	Jul 26, 2002
Digital Map	OpenStreetMap http://www.openstreetmap.org	Every Day
Geo-database	EarthExplorer (US. Geological Survey) http://earthexplorer.usgs.gov/	Feb 28, 2013
Geo-database	SRTM30 PLUS http://topex.ucsd.edu/WWW_html/srtm30_plus.html	Jan 16, 2013
DEM	Shuttle Radar Topography Mission (SRTM) http://www2.jpl.nasa.gov/srtm/cbanddataproducts.html	Aug 4, 2009
Geo-database	NGA GEOnet Names Server (GNS) http://earth-info.nga.mil/gns/html/	Apr 1, 2013
Maps	David Rumsey Map Collection http://www.davidrumsey.com/	Feb 11, 2013
Geo-data	NASA Visible Earth Project http://visibleearth.nasa.gov/	Apr 26, 2013
Geo-data	ESRI World Basemap Data http://www.esri.com/data/free-data	May 28, 2010
Geo-data	Harvard Map Collection http://hcl.harvard.edu/libraries/maps/	Unknown
Geo-database	Virtual Terrain Project http://vterrain.org/	Unknown
DEM	Global Multi-resolution Terrain Elevation Data 2010 https://lta.cr.usgs.gov/GMTED2010	Feb 28, 2013

Table 2: List of geo-software/packages with free or open source licenses. The top six (in italics) were used.

Software/package	Field of interest (if any specific)	Language	License
<i>Google Earth</i>	<i>3D visualisation</i>	<i>Unknown</i>	<i>Freeware</i>
<i>Quantum GIS</i>	<i>GIS</i>	<i>C++</i>	<i>GPL</i>
<i>GRASS</i>	<i>GIS</i>	<i>C, C++, Python</i>	<i>GPL</i>
<i>SAGA</i>	<i>GIS, Analysis</i>	<i>C++</i>	<i>Open source</i>
<i>Sextante GIS</i>	<i>Analysis</i>	<i>C</i>	<i>Open source</i>
<i>Orfeo Toolbox</i>	<i>Remote sensing analysis</i>	<i>C</i>	<i>Open source</i>
ClusterPy (non-GUI, open source)	Spatial clustering	Python	BSD ³ license
GeoGrouper (GUI, free)	Spatial clustering	Python	BSD license
PySAL	Analysis	Python	BSD license
Gmap Creator	Web mapping	Unknown	Freeware
Geo DaSpace	Spatial econometrics	Python	Freeware
SANET	Point patterns in networks	Unknown	Freeware (terms)
CrimeStat	Spatial analysis	Unknown	Freeware
SaTScan	Epidemiology	Unknown	Freeware
Croizat	Pan biogeography	Python	GPL
Hawths Tools	Ecology	Unknown	Freeware
Fragstats	Landscape metrics for categorical map patterns	Visual C++	Freeware
Whitebox GAT	GIS and remote sensing package	.NET framework 3.5	GPL
ILWIS Open	Remote sensing analysis	Unknown	GPL
Map Comparison Kit (MCK)	Categorical and numerical map comparisons	C++	Freeware
Dyn TM	Web mapping	Under development	Unknown
GeoDaNet	Point patterns in networks	Python	Under development

³ BSD license is a family of permissive free software licenses with minimal restrictions on the redistribution of software.

OSM data can be exported from its official website⁴ and the historical updates of the datasets are recorded and can be recovered. Large OSM datasets also can be downloaded from Planet.osm⁵, where the global OSM data are stored as an integrated file. Under the Open Data Commons Open Database License 1.0, a new version of the global OSM data is released weekly (currently every Thursday morning) with a file size of over 370 GB of uncompressed XML variant format or 27 GB for compressed data.

4. SATELLITE IMAGES ACQUISITION FROM GOOGLE MAPS

Google Maps is a web mapping service application with technology provided by Google. Similar to other map services providers, Google Maps powers many map-based services, including the Google Maps website, maps for mobiles, and maps embedded in third party websites by using the Google Maps API. Google has contracted with both government and private satellite owners to receive images so that every part of the world can be mapped at the same level of detail. Higher resolution images are used for more interesting locations and more populated places, such as urban areas. Rural areas tend to have less detail in the imagery. Currently, the best resolution is less than 60 cm per pixel and the worst is about 450 cm per pixel.

Google allows the use of maps and Earth satellite images for non-commercial, non-violent and non-invasive purposes only. Google obscures sensitive or private information and does not allow any commercial use of their images. Google satellite images are accessible by both API and internet browsers. Methods of acquiring satellite imagery from Google Maps with Firefox and its add-on screen grab tool were published by Bimpson (2010) on the DetailDesignOnline website⁶.

In our example, the study region is the Canberra urban area with latitude range [-35.150000, -35.501491] and longitude range [149.000000, 149.208516]. For our research, this whole area is further subdivided by a 20 (longitude) by 40 (latitude) grid system with a total of 800 grid cells. The size of each grid cell is 1 km × 1 km. To calculate the corresponding step changes in latitude and longitude on Google Maps, length variation in Canberra needs to be calculated.

According to U.S. Government's National Geospatial-Intelligence Agency (NGA), the step distance with latitude and longitude in Canberra can be calculated by Equations (1) and (2) respectively, using the mean latitude of -35.351491. Based on these equations, the length step of 1° of latitude is 110.9519 km and the length step of 1° of longitude is 90.193 km. Comparing these to the results 110.947 km and 90.896 km obtained from the NGA Calculator for a reference location within Canberra, our estimates are fit for purposes. The differences occur because we are using the mean latitude to standardize our calculation.

$$\Delta_{LAT}^1 = 110.852 + \frac{111.132 - 110.852}{45 - 30} \times (35.351491 - 30) = 110.9519 \text{ km} \quad (1)$$

$$\Delta_{LONG}^1 = 96.486 - \frac{96.486 - 78.847}{45 - 30} \times (35.351491 - 30) = 90.193 \text{ km} \quad (2)$$

With a fixed length of 1 km both horizontally and vertically, the step of each grid is 0.0109745 degrees in latitude and 0.0090126 degrees in longitude. With a monitor display resolution 1280×1024 pixels, which was used for Google image acquisition, the HTML "frame size" is 8200 pixels horizontally and 8300 pixels vertically. Using this method, a total of 800 Google satellite images (each 1 km × 1 km) of the Canberra urban area were acquired. The acquisition process is illustrated in Figure 1.

5. FEATURES DETECTION AND EXTRACTION

One of the aims of our research project is to enhance resistance to natural disasters. To protect both public and private properties, geo-informatics of houses, buildings, streets, vegetation and other land covers are required for vulnerability measurements and risk modelling. The high resolution Google satellite images acquired by the above method could not directly provide such information. Feature detection and extraction techniques were used to produce required geo-informatics of the entity of value.

To achieve a cost-free and open source geo-software, Orfeo Toolbox (OTB) was used for this task. OTB is a library for both optical and radar remote sensing image processing, which is implemented in C++ based on ITK⁷. The OTB project was initiated by CNES (French Space Agency) in 2006 and is under extensive development with participation from the open source community.

⁴ OpenStreetMap: <http://www.openstreetmap.org/>.

⁵ Planet.osm homepage at <http://planet.openstreetmap.org/>.

⁶ DetailDesignOnline website: <http://www.detaildesignonline.com>.

⁷ ITK (Insight Toolkit): <http://www.itk.org>.

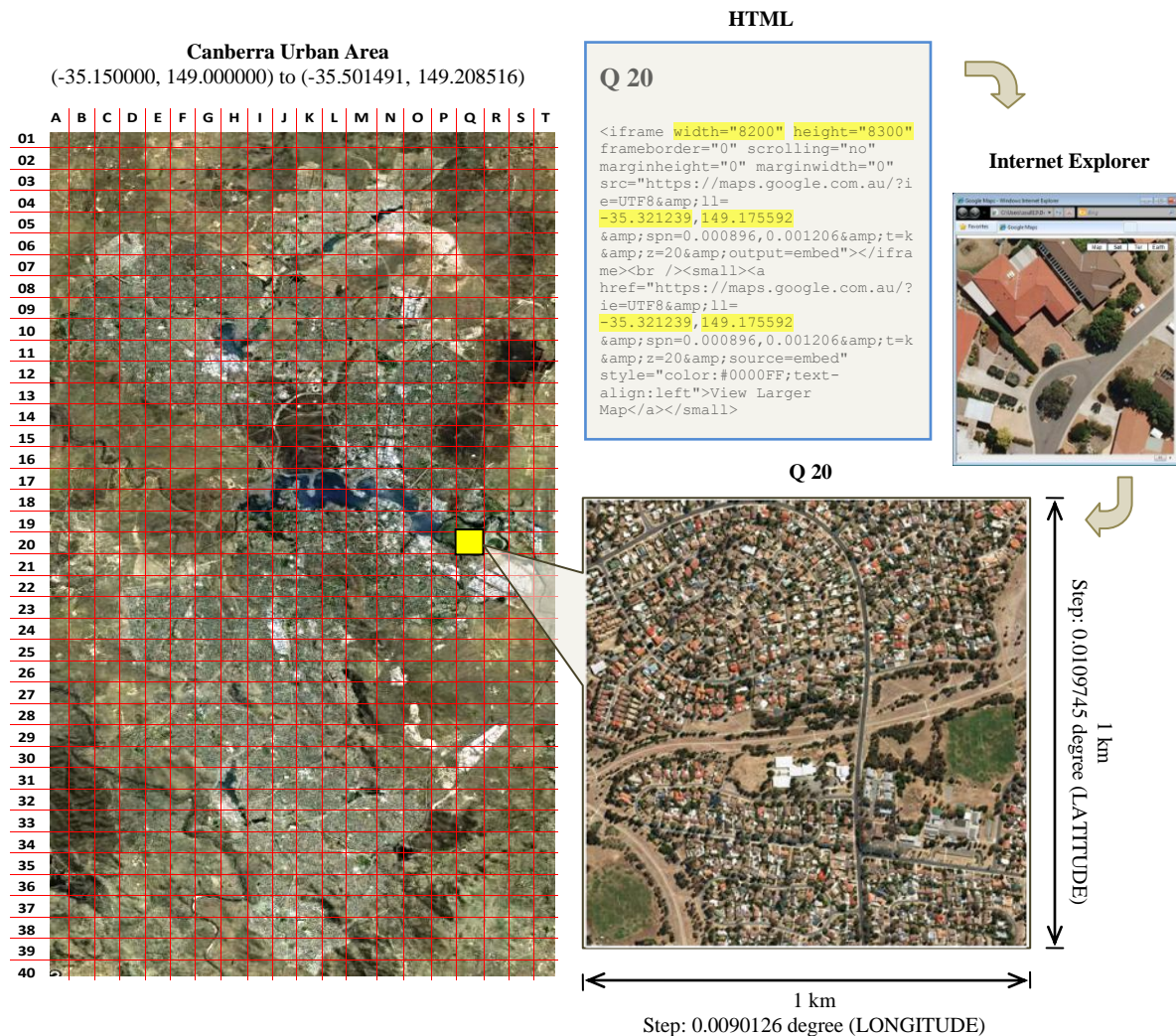


Figure 1. Satellite images acquisition from Google Maps using Internet Explorer.

5.1. Segmentation

Segmentation produces homogeneous clusters or objects, with each of which is associated additional attributes such as mean, variance, shape index, textural measures, etc. depending on the segmentation method used. The aim of segmentation is to reduce the complexity of the input data and to increase the accuracy of the classification (see the next subsection). The level of details available in high resolution images can have a strong negative effect at several stages of processing. For instance, the details of roof superstructures are irrelevant for the purpose of extracting the whole building footprint. The mean-shift algorithm provides us an efficient way to simplify such images (Inglada and Christophe, 2009). In addition, mean-shift segmentation also helps us to separate the original imagery into multi-scale spaces allowing the incorporation of scales into classification (Vu, 2012). A comparison of an original image and the result from the OTB mean-shift segmentation (with parameters settings: spatial radius = 30, spectral value = 45.5, and minimum region size = 100) is shown in (a) and (b) of Figure 2.

5.2. Support Vector Machines (SVM) Classification

The OTB SVM classification provides a supervised pixel-based classification chain based on learning from multiple images. It supports large images through streaming and multi-threading (OTB Development Team, 2012). The classification chain performs the SVM training steps based on the intensities of each pixel as features. A standard OTB SVM classification chain has four steps: (1) image statistics estimation; (2) building up a training dataset; (3) training the SVM model; and (4) SVM classification. The SVM model training should be done with local images.

The trained SVM model was used as an input parameter in OTB application of ‘ImageSVMClassifier’ to perform the SVM classification using the ‘decision rule’ which was learned from the training dataset. The results of an SVM classification is a labeled raster file. The SVM labelled results and the color mapped results are shown in (c) and (d) of Figure 2. In the color mapped results, red represents roofs, green represents vegetation, yellow represents land, and dark blue represents roads.

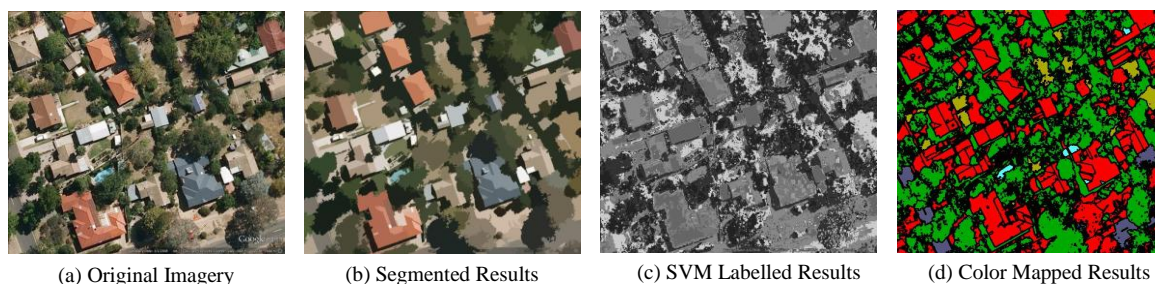


Figure 2. Comparison of original imagery, segmented results, SVM labelled results and colour mapped results.

Due to limitations of the original data sources and the acquisition methods, elevations of buildings/houses cannot be extracted from the captured Google satellite imagery. Elevations also cannot be estimated by the produced geo-informatics. It should be pointed out that, although a building’s elevation may be estimated by its shadow length on satellite image, we found it is difficult to measure the shadow length. This is because trees and houses in an urban environment are usually too close to each other; their shadows are often either mixed together or are projected on walls. Hence, the shadow length cannot be used to estimate the elevation in this case. Individual building heights may be estimated by Google Building Maker and Google SketchUp manually.

6. ATTRIBUTES ENRICHMENT

Particular attributes of features related to location information (such as buildings’ elevations) and disaster characteristics are required for vulnerability modelling; however they are missing in the geo-informatics extracted from satellite images. There is a solution to the problem by merging with other geographic and demographic information so that the geo-informatics attributes can be enriched. In this section, three methods are introduced to illustrate how to enrich attributes from raster layers, vector layers, and Microsoft Excel files.

A core QGIS plugin tool called “Zonal Statistics” can be used to set up mean digital elevation values to features according to their locations. Extracted feature attributes can be enriched with information carried by all raster files. Zonal Statistics has several predesigned statistics such as the count, sum and mean. More statistics using complex calculations can be produced by the “Grid Statistics” algorithm from SAGA GIS, which is also accessible from the Sextante toolbox.

Attributes can also be enriched from vector layers, such as the OSM layers and other shape files. The OSM vector layers and vector layers from other reliable sources (e.g., Geoscience Australia website: <http://www.ga.gov.au/>) contain a lot of useful information which can be used to enrich the produced geo-informatics. QGIS function “Join Attributes by Location” can combine attributes of two vector layers based on their locations. The produced vector layers can also be enriched by combining them with Microsoft Excel files or csv files, which must have columns of x and y coordinates. To input and output Microsoft Excel tables, the python libraries of python-xlrd and python-xlwt are required.

Attributes of the produced geo-informatics can also be enriched from multiple sources of data by converting them into raster layers, vector layers and Excel files. More comprehensive attributes of the produced geo-informatics can be further enriched from all available data using spatial analysis packages in the language R.

7. CONCLUSIONS

Our experience with practical approaches to zero-cost collection, acquisition and production of geo-informatics has been presented. This is a part of a current research project for urban vulnerability mapping and quantitative risk assessment of natural disasters. An abundance of internet-accessible free geo-data and open source geo-software were listed. Timely high resolution satellite images can be acquired from Google Maps and Google Earth using several methods. In this paper an internet browser based acquisition method was introduced. Orfeo Toolbox as a free and open source library for remote sensing tasks is useful and

effective for satellite imagery analysis and features extraction. Attributes of the produced geo-data can be enriched from other data sources using QGIS. Although there are limited features compared to the commercial geo-data and geo-software resources, free geo-data and open source geo-software have the advantages of allowing further developments by users to meet practical needs with flexibility. The detailed approach can be found in Zou and Lin (2013).

ACKNOWLEDGMENTS

The research in “Urban Natural Disasters Risk Assessment and Community Vulnerability Measurement Mapping” was supported by both the Chinese Scholarship Council and the Australian Commonwealth Scientific and Industrial Research Organisation (CSIRO). The authors thank Dr. Warren Jin (CSIRO, Division of Computational Informatics) for his administration of QGIS on a super computer, which was used to produce the geo-informatics presented in this paper.

REFERENCES

- Bennett, J. (2010). *OpenStreetMap*. Packt Publishing Ltd.
- Bimpson, A. (2010). Capturing high resolution satellite imagery from google maps. Retrieved on 28 September 2013 from <http://www.detaildesignonline.com/software/1849>.
- Christophe, E. and Inglada, J. (2009). Object counting in high resolution remote sensing images with OTB. *Geoscience and Remote Sensing Symposium, 2009 IEEE International, IGARSS 2009*, 737–740.
- Coast, S. (2011). How OpenStreetMap is changing the world. In K. Tanaka, P. Fröhlich and K.-S. Kim (Eds.), *Web and Wireless Geographical Information Systems SE – 2, Vol. 6574*, p4. Springer Berlin Heidelberg.
- Girres, J.-F. and Touya, G. (2010). Quality assessment of the French OpenStreetMap dataset. *Transactions in GIS, 14*(4), 435–459.
- Haklay, M. and Weber, P. (2008). OpenStreetMap: user-generated street maps. *IEEE Pervasive Computing*, 7(4), 12–18.
- ITN. (2012). ESRI donation supports collaboration with OpenStreetMap community. *Information Technology Newsweekly*, 73.
- Lin, Y.W. (2011). A qualitative enquiry into OpenStreetMap making. *The New Review of Hypermedia and Multimedia*, 17(1), 53–71.
- OTB Development Team (2012). *The Orfeo ToolBox Cookbook: a guide for non-developers updated for OTB-3.14*. Centre National D’etudes Spatiales, France.
- PR Newswire (2012). ESRI donation supports collaboration with OpenStreetMap community: ArcGIS users continue to contribute to the OSM initiative. *PR Newswire*.
- UNISDR. (2012). Statistics of Natural Disasters (1980-2011). Retrieved on 28 September 2013 from <http://www.unisdr.org/we/inform/disaster-statistics>.
- Vu, T. (2012). Object-based remote sensing image analysis with OSGeo tools. In *Applied Geoinformatics for Society and Environment, FOSS4G Southeast Asia, 5th International Summer School and Conference*, 79–84.
- Ying, F., Mooney, P., Corcoran, P. and Winstanley, A. (2010). Polygon processing on OpenStreetMap xml data. In *Proceedings of the GIS Research UK 18th Annual Conference, University College London, London, England*, 149–154.
- Zou, Z. and Lin, X. (2013). Geo-informatics Production for Urban Disasters Risk Reduction: a Zero-Cost Solution. In 2013 International Conference on Geo-Informatics in Resource Management & Sustainable Ecosystem (GRMSE2013), Wuhan, China, 8-10 November.
- Zschocke, T. and León, J. (2010). Towards an ontology for the description of learning resources on disaster risk reduction. In M. Lytras, P. Ordonez De Pablos, A. Ziderman, A. Roulstone, H. Maurer and J. Imber (Eds.), *Knowledge Management, Information Systems, E-Learning, and Sustainability Research SE – 8 Vol. 111*, 60–74. Springer Berlin Heidelberg.