Utilizing Satellite Optical Imagery for Building Footprint Extraction and City Modelling

Haval Sadeq, Jane Drummond, Zhenhong Li

School of Geographical and Earth Sciences,
University of Glasgow,
Glasgow G12 8QQ, United Kingdom
Telephone +44(0)1413301687
Fax: +44(0)1413304894
Email of corresponding author h.sadeq.1@research.gla.ac.uk
www.glasgow.ac.uk

1. Introduction

3D models provide information for tourists, planners, the emergency services, the military, geographers, etc. For decades 3D modelling has relied on aerial photography or more recently laser based technology (LiDAR and terrestrial lasers canning) but these may be expensive and are sometimes difficult to acquire due to security issues. Therefore the challenge is to use an alternative data source, which is cheaper and more accessible, such as satellite imagery. Satellite imagery is available anywhere, and at lower cost per unit area when compared to photography or laser based sources.

This research focuses on using high resolution satellite imagery as an alternative to Aerial photography or LiDAR data, for automatically building footprint extraction and 3D modelling. The research has extensively utilized Bayesian theory to optimize the boundary of the building footprint and find the most probable building corners. The level of building detail hoped for is LoD1 or Level of Detail 1 (Meng et al. 2007); a level which can be achieved by extruding the building footprint and assuming the roof of the buildings are flat and reach a maximum height. The method applied can be summarised as four main stages:

- data preparation which consist from producing and preparing the DSM and orthophotography from satellite imagery;
- building detection, in this stage the building are identified and used as input in Bayesian theory;
- building extraction which consists from identifying the location of the building, and finally
- building construction and 3D model construction.

The structure of the paper is as follows: data preparation is discussed in section 2; building detection is illustrated in section 3. Building extraction is explained in section 4. Building construction, Bayesian regularization and 3D modelling are shown in section 5. Finally the experimental result and assessment is discussed in section 6.

2. Data Preparation

Two pairs of high resolution satellite images were used in this study. The images that were acquired were of Glasgow, from two different sources - World View01 and Pleiades. Both have spatial resolution GSD equal to 0.5m as shown in Figure. 1.

(a) (b)

Figure 1. High resolution satellite imagery (a) World View satellite imagery GSD 0.5m, (b) Pleiades satellite imagery GSD 0.5m.

The characteristics of the images are shown in Table 1; the World View images were panchromatic whereas the Pleiades images were pansharpened. The Pleiades imagery has been produced from fusing multispectral low resolution imagery of 2m spatial resolution with panchromatic imagery of 0.5m resolution.

Table 1. World View01 and Pleiades along track satellite imagery properties

Image ID	Acquisition Date	Max off Nadir(deg)	Sun Elevation(deg)	Image Type	Band No.	Res.GSD (m)
WV01 image1	24/05/2012	17.33	54.59	Panchromatic	1	0.5
WV01 image2	24/05/2012	21.01	54.63	Panchromatic	1	0.5
Pleiades image1	09/07/2013	11.91	55.46	Pansharpened	4	0.5
Pleiades image2	09/07/2013	14.03	55.51	Pansharpened	4	0.5

2.1 DSM and Orthophotography Generation

The digital surface model (DSM) and Orthophotography were generated using Socet GXP software. Ground control points were obtained using precise GPS, and have been used in the triangulation. The resolution of the orthophoto has been produced with a GSD equal to 0.5m. The DSM resolution was initially 1.5 metres, but was later re-sampled to the same resolution as the orthophotography, namely 0.5m.

2.2 Producing, segmenting and labelling nDSM

The DSM was further processed to produce a so-called 'nDSM' or normalised digital surface model, to be used for building detection. The traditional procedure to obtain an nDSM is by subtracting the DTM (digital terrain, or bare-ground, model) from the DSM. But it was difficult to get a DTM with equivalent or similar accuracy to the DSM, therefore another method has been implemented to produce the nDSM - using a mathematical morphology image processing technique. Later, the resulting nDSM was thresholded in order to eliminate low ground surface objects, such as bushes, while retaining above ground objects such as individual buildings. The final stage in producing the nDSM involves assigning a label to each building using the connected component labelling algorithm, see Figure 2(a).

(a) (b) (c)

Figure (2) (a) the labelled nDSM for the study area; (b) result of applying the Minimum Error Threshold on the orthophotography; (c) result of applying the Inter-Modes Threshold on the orthophotography.

2.3 Segmenting the Orthophotography and Producing the Edge Map

The nDSM, whose production is summarised in section 2.2, represents buildings as irregular blob shapes. Improvements were necessary and the orthophotography was used for this. First the orthophotography was enhanced by sharpening and applying a mean shift algorithm, then thresholding techniques were used to classify the orthophotography into segmented building, tree and other areas. The thresholding method first used was the Minimum Error Algorithm (Kittler & Illingworth, 1986), but this threshold method was not able to detect high reflectance buildings (e.g. corrugated aluminium sheeting), therefore another method of thresholding based on the Inter-Modes Algorithm (Prewitt & Mendelsohn 1966) was implemented, see Figure 2(b) and 2(c). As well as the segmented orthophotography, an 'edge map' was produced by implementing Canny Edge Detection using the enhanced orthophotography. The edge map is produced in order to use it in regularizing building boundary within an application of Bayesian theory, as discussed in section 6. Some unnecessary details in the edge map, particularly on the roofs of the buildings have been

removed using mathematical morphology. Some building edges are not detected in the edge map due to low contrast in some parts of the orthophotography. The edge map was subsequently further improved, also using mathematical morphology, leading to the 'modified edge map'.

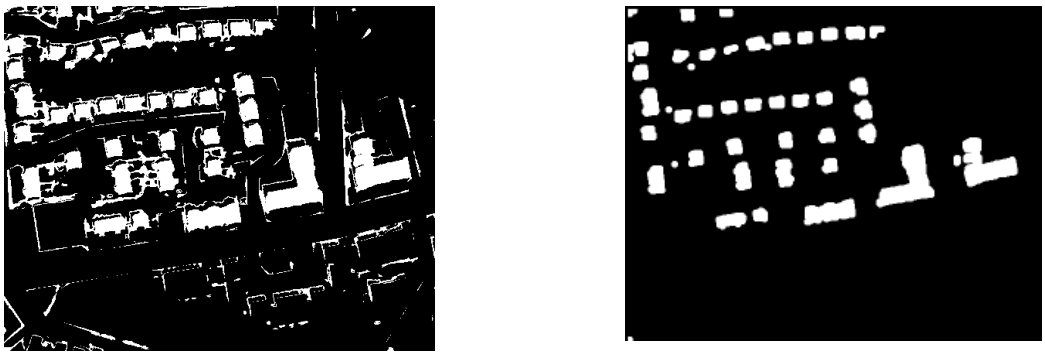
2.4 Road, Shadow and Green Area Masking

In the segmented orthophotography, the roads were detected together with the buildings, consequently causing problems. It was necessary to filter out the roads from the image, keeping the buildings only. Much research has been conducted in road detection, but, not being the focus of this research and for the time being, the roads were extracted manually using ESRI's ArcGIS software.

The shadows also were a problem in detecting buildings, because they were retained as a part of their building, therefore it was necessary to identify and eliminate them. The principle used to detect shadow assumes they have a lower intensity value than other objects, and has been used by Yamazaki et al. (2009) and Liu et al. (2011). The detected shadows are fragmented, therefore mathematical morphology has been applied to convert them to regions.

As mentioned, the nDSM has been used to detect buildings, with their distinctive elevation, from the segmented image; since trees have elevation they, too, have been detected in the nDSM as buildings. For that reason an NDVI data set has been produced and used for masking out the green area. The data source used for this purpose was Pleiades, because its four bands included both red and infra-red.

After detecting the roads, shadows and green areas, all were subtracted from the thresholded segmented orthophotography, thus the remaining objects in the binary image were only the buildings as shown in Figure 3(a). Finally applying the mathematical morphology operator to separate the objects from each other and removing the noise, kept the bare buildings (Figure 3(b)).



(a) (b)

Figure 3 thresholded binary images (a) the result of the images after subtracting the roads, shadows and green areas; (b) using mathematical morphology in opening and closing the binary image.

3. Building Detection

The labelled nDSM has been used to detect the individual buildings. Buildings in the nDSM were detected as 'blobs', almost always representing just part of a building. It was thus necessary to expand each blob using mathematical morphology's dilation operation. Later

each expanded blob's location was found in the binary image and the corresponding building identified. These identified buildings were used as rudimentary initial building locations in the next stage.

4. Building Extraction

At this stage the buildings were represented as pixels. In order to regularize the boundary it was necessary to vectorise them. First pixel locations were converted into local national grid WGS84 co-ordinates, later on points were sorted using the Nearest Neighbour Algorithm (NNA) because they had been randomly sorted. The boundaries were now represented by multiple points, even straight lines, and consisted of many segments. It was difficult to use the boundary, at this stage, because it contained many corner points. The number of points was thus decreased only keeping significant corner points; the method used was that of Ramer-Douglas-Poiker. At this stage the building boundary obtained could be used as an initial building boundary for the eventual final estimation of the building footprint location.

5. Building Construction, Regularization and 3D modelling

Albeit the initial buildings boundary has been produced from the segmented image, they do not represent the actual building location. They have been modified during the refinement of the segmented image using mathematical morphology, which can lead to a loss of accuracy, moreover the boundary was not regular. In order to increase the accuracy of the initial building boundary and improve the buildings shape, Bayesian theory has been applied.

In order to use Bayesian theory it was necessary to identify the data set most suited to be the measurement base. For that purpose an edge map has been used. The model that was applied at this stage was originally developed by Wang et al. (2006) to extract building boundaries from LiDAR data. It is based on finding the most probable coordinates of the building footprints, based on the 'modified edge map'. Bayesian theory depends on the prior information. The prior information used in this study was that building corners are 90^o and the building edges are 180^o. These assumptions are visualised in Figure 4, where the 'penalties' required by the implemented Bayesian approach are seen to be the lowest for 'corners' of 90 or 180 degrees.

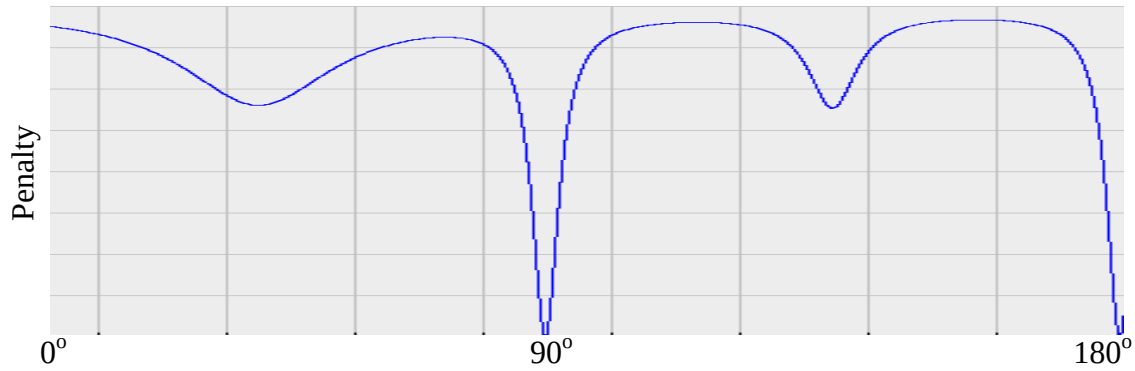


Figure 4. Prior probability represented by functionf (α) is applied to regularize a boundary, showing that 90 and 180 degrees has the lowest penalty than other angle values.

Bayesian theory generates an outcome probability (posterior probability) based on an initial probability (prior probability: p(Θ)) and a likelihood (p(X| Θ)) – see Equation 1.

$$p(\text{buildgin footprint}|\text{location of boundary points}) = p(\Theta|X) = \frac{p(\Theta)p(X|\Theta)}{p(X)} \quad --(1)$$

The prior function can be defined by using the above graph, and represented by $f(\alpha_i)$:

$$P(\Theta) \propto \prod_{i=1}^{k} f(\alpha_i) \quad ------(2)$$

Where k is the number of angles in the building footprint,

The likelihood function can be defined using error values (i.e. σ), assuming errors are normally distributed

$$P(X|X\Theta) \propto \prod_{i=1}^{n} \left\{ \frac{1}{\sqrt{2\pi\sigma}} EXP\left(-\frac{d^2}{2\sigma^2}\right) \right\} --- (3)$$

The normalization factor does not need to be calculated because our aim is to maximize the functions

$$\bar{\Theta} = \text{argmax}\{P(\Theta).P(X|\Theta)\} = \text{argmin}\{-\log\{P(\Theta).P(X|\Theta)\}\} \text{-----(4)}$$

The final optimization model for the building regularization will be:

$$\bar{\Theta} = \text{argmin}\left\{ -\sum_{i=1}^{k} \log f(\alpha_i) + \frac{1}{2\sigma^2}\sum_{j=1}^{n} d_j^2 \right\} \ldots (5)$$

By minimizing the equation we will get the result for the boundary as shown in figure 5(c).

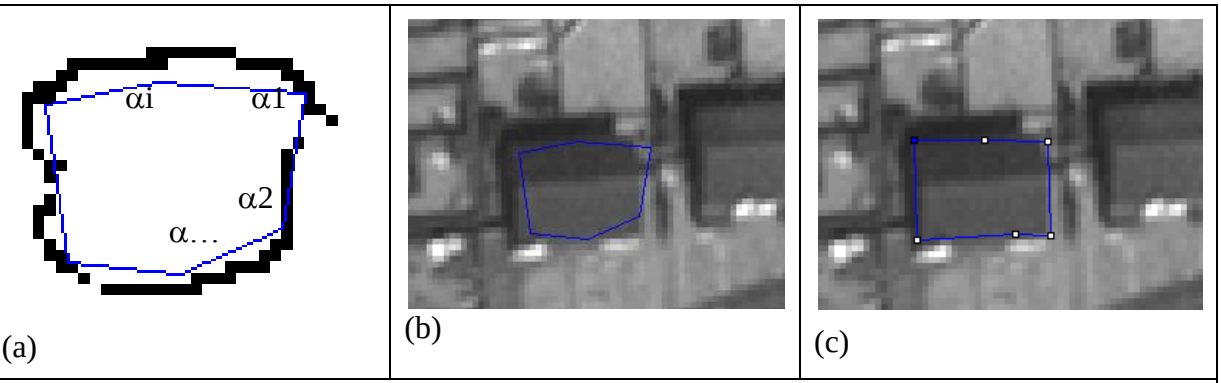


Figure 5 (a) initial building boundary overlaid over edge map, showing the corner angle; (b) initial building boundary that is obtained from segmented image which is been used in Bayesian regularization; (c) result of regularized building boundary which is obtained by minimizing equation 5.

A resulting building footprint is shown in Figure 6(a, b, c). In addition to the building footprint construction, a 3D model has also been produced. This is constructed by extracting the elevation of the buildings from the DSM and combining it with the regularized building boundary. The level of the detail which is produced in this research is LoD1 because it uses an extruded building footprint. The extrusion is based on the height which is obtained from DSM by taking the maximum height in within the building footprint.

Figure 6(b) and (c) shows the comparison of DSM produced from Socet GXP and the result of applying our algorithm

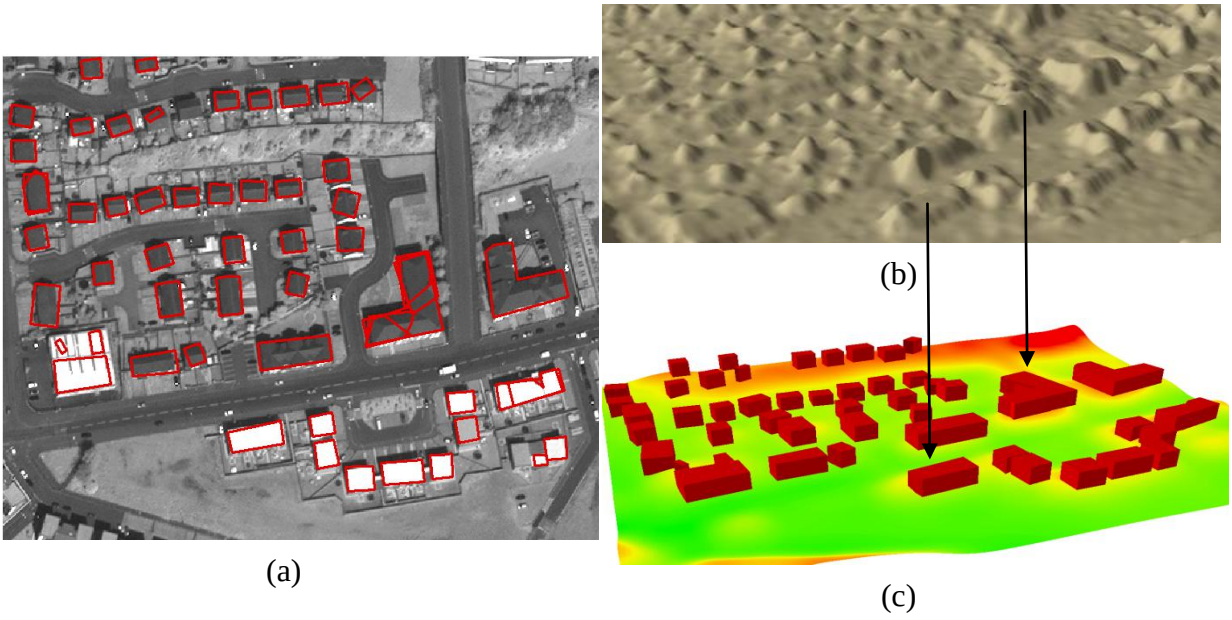


Figure 6 result of Applying Bayesian theory (a) Regularized building boundary overlaid on the orthophotography; (b) the produced DSM from Socet GXP showing the building; (c) 3D model of the study area.

6. Results and Discussion

The applied method for building footprint extraction and 3D modelling construction from high resolution satellite imagery was fully automatic except for the road extraction. The proposed method has provided reliable results without depending on any other data as reference, such as a 2D cadastre database, as mentioned in Tack et al. (2012). The result of the applied method has shown that the extraction percentage is about 96%, among 51 buildings only 2 failed to be detected. It is worth mention that 90% of the 49 are extracted without apparent rotation; five of the buildings, which is equivalent to 10%, have undergone some slight rotation. The result of the applied algorithm showed it is better than the algorithm applied by (Aytekin et al. 2009) which gave result of building extraction of about 81%. Another study by (Dahiya et al. 2013) showed that among 122 building, 24 are not identified at all and 18 wrongly identified.

7. References

AYTEKIN, O., ULUSOY, I. & ERENER, A., 2009. Automatic and unsupervised building extraction in complex urban environments from multi spectral satellite imagery. Recent Advances in Space Technologies, 2009. RAST '09. 4th International Conference on, pp.287–291.

DAHIYA, S., GARG, P.K. & JAT, M.K., 2013. Building Extraction from High Resolution Satellite Images. International Journal of Computing Science and Communication Technologies, 5(2).

KITTLER, J. & ILLINGWORTH, J., 1986. Minimum Error Thresholding. Pattern Recognition, 19(1), pp.41–47.

LIU, J., FANG, T. & LI, D., 2011. Shadow Detection in Remotely Sensed Images Based on Self-Adaptive Feature Selection. IEEE Transactions on Geoscience and Remote Sensing, 49(12), pp.5092–5103.

MENG, L., FORBERG, A. & SARJAKOSKI, T., 2007. 3D Building Generalisation. In Challenges in the Portrayal of Geographic Information: Issues of Generalisation and Multi Scale Representation. Elsevier Science Ltd., pp. Elsevier Science Ltd., Oxford, pp. 211–232.

PREWITT, J.M. & MENDELSOHN, M.L., 1966. The analysis of cell images. Annals of the New York Academy of Sciences, 128(3), pp.1035–53.

TACK, F., BUYUKSALIH, G. & GOOSSENS, R., 2012. 3D building reconstruction based on given ground plan information and surface models extracted from spaceborne imagery. ISPRS Journal of Photogrammetry and Remote Sensing, 67, pp.52–64.

WANG, O., LODHA, S.K. & HELMBOLD, D.P., 2006. A Bayesian Approach to Building Footprint Extraction from Aerial LIDAR Data. Symposium A Quarterly Journal In Modern Foreign Literatures, pp.0–7.

YAMAZAKI, F., LIU, W. & TAKASAKI, M., 2009. Characteristics of shadow and removal of its effects for remote sensing imagery. In 2009 IEEE International Geoscience and Remote Sensing Symposium. IEEE, pp. IV–426–IV–429.

Biography

Haval Sadeq: received a BSc in Civil Engineering in 1996 and an MSc in photogrammetry in 2000 from College of Engineering/Salahaddin University-Hawler. In 2011 he started his PhD (still in progress) at the University of Glasgow. His research interest is in photogrammetry, remote sensing, image processing and Bayesian statistics.

Dr Jane Drummond and Dr Zhenhong Li are, currently, Senior Lecturers at the University of Glasgow, and are Haval's PhD supervisors.