Snake-based approach for building extraction from high-resolution satellite images and height data in urban areas

Tao Guo*
Yoshifumi Yasuoka**

*Institute of Industrial Science, University of Tokyo
Ce-509, 4-6-1 Komaba, Meguro, Tokyo 153-8505, Japan
Email: guotao@iis.u-tokyo.ac.jp,
Tel:+81-3-5452-6415, fax:+81-3-5452-6410

**Institute of Industrial Science, University of Tokyo
Ce-509, 4-6-1 Komaba, Meguro, Tokyo 153-8505, Japan
Email: yyasuoka@iis.u-tokyo.ac.jp,
Tel:+81-3-5452-6409, fax:+81-3-5452-6480

ABSTRACT

In urban areas, it has been long a challenge topic to automatically extract urban objects from images due to the high object density and scene complexity. Applying normal image processing methods could not achieve satisfied performance, especially for high-resolution satellite images. This paper presents a snake-based approach for building extraction from high-resolution IKONOS satellite images by combining height data. Multiple cues derived from both data sources are integrated into our snake model to precisely control the movement of contours so that the boundary of buildings can be located with good accuracy and efficiency. Some experiment results are presented as well in this paper.

Keywords: Building extraction, DSM, IKONOS, snake models

1. INTRODUCTION

The 3D city model has been an increasing interest for recent several years because its significant role in many application areas, like communication industry, urban microclimate and pollution control analysis, transportation navigation, landscape planning and visualization, etc. One of essential issues for the 3D city models generation is building extraction. Survey and photogrammetry are traditionally ways to map details in urban areas even though it is well known that they are resource and time consuming. Recently, high-resolution satellite imagery has become widely available, and this shows an opportunity for so-called urban remote sensing to challenge the topic of urban details mapping.

However, in urban areas, frequently it is very tough to extract target objects solely based on images in the full- or semi- automatic approaches due to high object density and scene complexity. Therefore, it is an important aspect to combine other clues for this purpose. Height data has been approved very valuable information for raised objects discrimination. Also as the third dimension, height is required for 3D models. Height data can be acquired by automatic image matching algorithms applying stereo aerial or high-resolution satellite imagery or can be directly captured by airborne laser scanning system. But in urban areas, stereo matching method has a big problem due to occlusions and height discontinuities. Therefore, the direct height measurement by airborne laser scanners usually provides digital surface models (DSM) data of higher and more homogeneous quality especially urban built-up areas (Haala, 1999).

A number of researchers have shown photogrammetric approaches for extracting building models from imagery, mainly aerial images/image pairs (Lang and Forstner, 1996, Henricsson, 1996, Gruen et al., 1997). Fraser, Baltsavias and Gruen (2001) reconstruct 3D building from high-resolution IKONOS stereo imagery. On the other hand, due to its advantages as an active technique for reliable 3D determination, airborne laser scanning (LIDAR) data has become a rather important information source for generating high quality DSMs. Weidner and Forntner (1995), Maas and Vosselman develop approaches for extracting building from DSMs or laser altimetry data. Meanwhile, integrating multiple data sources has been approved an efficient approach. Haala (1999) combines DSMs with color aerial images, Vosselman and Suveg (2001) fuse GIS maps with laser data and images show some good performances.

This paper describes our approach to extract 2D building outlines from high-resolution IKONOS satellite images and height data captured by airborne laser scanning system. Our data fusion scheme shows an efficient way to exploit the complementary characteristics of these two kinds of dataset for the purpos e of building extraction. The pyramid processing strategy is generally applied to extract features of objects from coarsely to finely. An adaptive snake-based approach is developed to cooperate multiple clues derived at various processing levels. The framework of our approach contains the following steps: (1) segmentation of raised objects from ground to generate normalized digital surface model (nDSM), (2) orthorectify images and derive edge etc. clues, (3) building blob detection by multiple height bins (MHB), and (4) building contour detection using snake model. This paper is structured as follows. Section 2 introduces the methodology and related models. Section 3 briefly describes the experiments. Finally, section 4 gives the conclusion and future work.

2. METHODOLOGY

Fig.1 illuminates the framework of our building extraction scheme. In order to simplify the complexity of processing, our approach generally follows top to down pyramid strategy. That is on higher level, emphasis is put on those global and apparent features, and then uses its output combining other information to refine details.

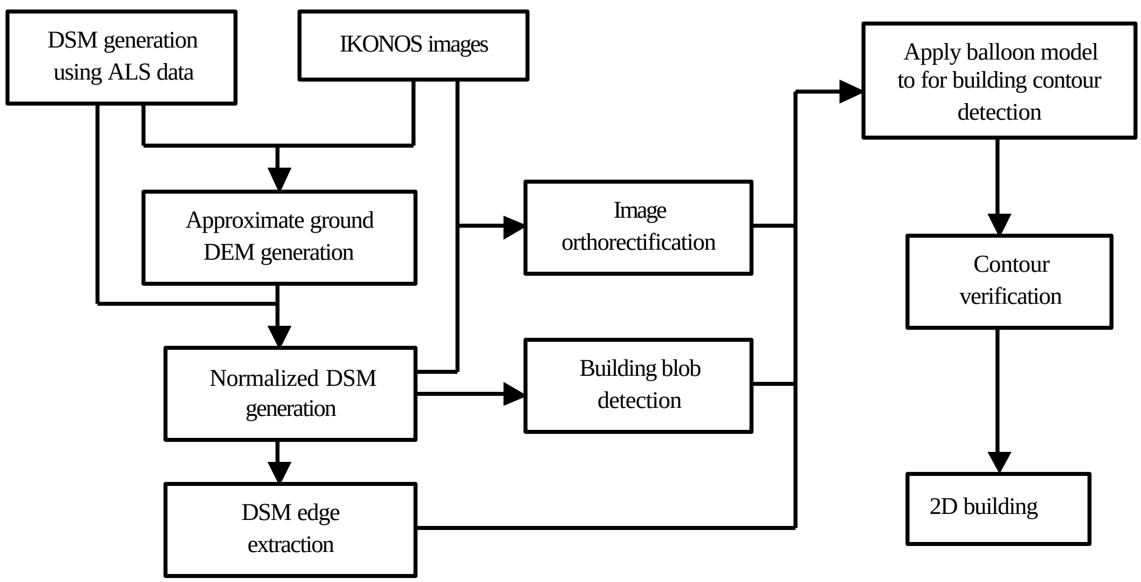


Fig. 1: Framework of building extraction scheme.

2.1 DSM and normalized DSM (nDSM) generation

Digital surface models (DSM) can be generated directly from interpolation of airborne laser scanning data, linear or higher order degree of surface curvature smooth edges of objects and lead to difficulties in discrimination buildings and trees. Therefore Nearest Neighbor (NN) method is used to interpolate ALS data into the same ground sampling distance (GSD) with IKONOS images in order to co-register with images, though it causes shifts in objects boundaries, shapes and edges of objects can be preserved, and also when grid size becomes small, this shift will decrease. DSM contains both terrain surface and objects above the terrain. In order to reduce the influence of slope of terrain, ground segmentation should be conducted. We have proposed an approach to combine the spectral signature taken from satellite images and height information of ALS data to coarsely generate an approximate ground DEM before apply morphological “opening” operation (Tao, G., Yasuoka, Y., 2001). In most cases, it will give good results. Subtraction of DTM from original DSM will generate the so-called normalized DSM, which results in a representation of objects rising from the terrain approximately put on a plane (Weidner and Forstner, 1995, Haala and Brenner, 1999).

2.2 Image orthorectification and edge information deriving

In order to directly co-register with DSM, images should be othorectified. But the orthoimages are produced from a DTM, and this cannot remove the vision difference caused by the looking angle of satellite sensor, particularly when the object is very high and satellite looking angle is not small enough, like in the downtown of City, such shift of high building sometimes is very big and has to be taken into account. For our case, the IKONOS images have been orthorectified using DTM, but raised objects have to orthorectified again using nDSM which is created as above. Ideally, rectified building should exactly match with the ALS data, but due to the noise and error of nDSM, practically still there is a deformation and displacement. More practical method is set a threshold, only when height is bigger than the threshold will be orthorectified, others low objects, just ignore the difference.

Edges information is very valuable clue of indicating existence of target object though it is normally not enough to represent the target features. In our scheme, edge information will be served as constraint clues for snake models. Both image gradient and DSM edge are derived in a similar way. For example, for deriving DSM edge, simply treating this nDSM as an image, a lot of edge detectors can be applied, e.g. Canny edge detector, Sobel detector to retrieve the height gradient information.

2.6 Snake models and building extraction

Snakes are also known as "active contour models " or deformable energy-minimizing curves, is introduced by Kass firstly. In the theory of snakes, the control points on the closed curve are guided by energy, which is defined as Eq. (1):

$$E = E_{\text{int}} + E_{ext} \tag{1}$$

where E_{int} is the internal energy formed by the snake configuration, E_{ext} is the external energy formed by external forces affecting the snake. The aim of the snake is to find a location that minimizes total energy.

$$E_{\text{int}} = E_{cont} + E_{curv} \tag{2}$$

E_{int} is the sum of the contour geometric constraints, defined as Eq. (2), where E_{cont} is the contour continuity energy. Minimizing E_{cont} over all the snake control points causes the snake control points become more equidistant. E_{curv} is the contour curvature energy. The smoother the contour is, the less is the curvature energy.

In Eq.(1), E_{ext} is the external energy and can be defined as Eq.(3):

$$E_{ext} = E_{img} + E_{con} \tag{3}$$

E_{img} is the image energy, which can be the image intensity or intensity gradient, and E_{con} is external constraint, variable constraints can be introduced into snakes model through this item. For each control point on snake contour, its total energy can be represented as Eq.(4) :

$$E_i = a_i E_{cont,i} + b_i E_{curv,i} + g_i E_{img,i} + l_i E_{con,i} \tag{4}$$

Where a , b , g , l are the weights of every kind of energy. The full snake energy is the sum of E_i over all the control points.

However, original snake model have its own limitations: firstly, if the initial contour or curve is not close enough to an edge, it is not attracted by it; secondly, if the initial contour is not submitted to any forces, it shrinks on itself; thirdly, it cannot handle concave/convex contours.

Cohen develops a so-called "balloon" model by adding a normal force to pull or push a snake contour towards its target object. If the force is put inside of snake contour, it makes snake contour to expand as a balloon until it reaches equilibrium against its internal resistance. In our schemes, we adopt the balloon model and develop the inflation force as follows:

$$F_i = k_1 \vec{n}(i) - k_2 \frac{E_{dsm}(i)}{|E_{dsm}(i)|} - k_3 \frac{E_{img}(i)}{|E_{img}(i)|} \quad \text{if } F_i > F_{threshold} \quad \text{else} \quad F_i = 0 \tag{5}$$

Where, F_i is inflation force, $\vec{n}$ is the normal unit force, as illustrated in Fig. 3. E_{dsm} is the energy generated from DSM edge and E_{img} is the energy derived from image gradient. Parameter k_1 is the movement amplitude of inflation force; k_2 and k_3 control the inflation force to sensitive to DSM edge and image gradient. $F_{threshold}$ is the threshold to adjust the influence of inflation force.

However, under the inflation force control points might be move too close or too far away from each other, and also it is not enough to accurately represent complicated shape of boundary by fixed number of control points, so a dynamic linear interpolation is developed to keep the distance between each point more constant. In the each iteration process, when the distance of adjacent points is less than threshold, then delete one control point among them, vice versa, when the distance is bigger than another threshold, a new control point is added at the middle position.

As illustrated above, a snake model needs an initial contour. In most cases, a manual process has to apply to make this initial contour. But our schemes aim at automatic processing, so we apply a process so-called multiple height bins (MHB) and it leads to what is referred to here as “elevation blobs”. Using MHB method, the nDSM heights are grouped into consecutive bins (height ranges) of a certain size. This results in segmentation of the nDSM in relatively few regions that are always closed and easy to extract. The MHB method is simple and fast and can be applied globally or locally (Baltsavias, E., Mason, S., and Stallmann, D., 1995).

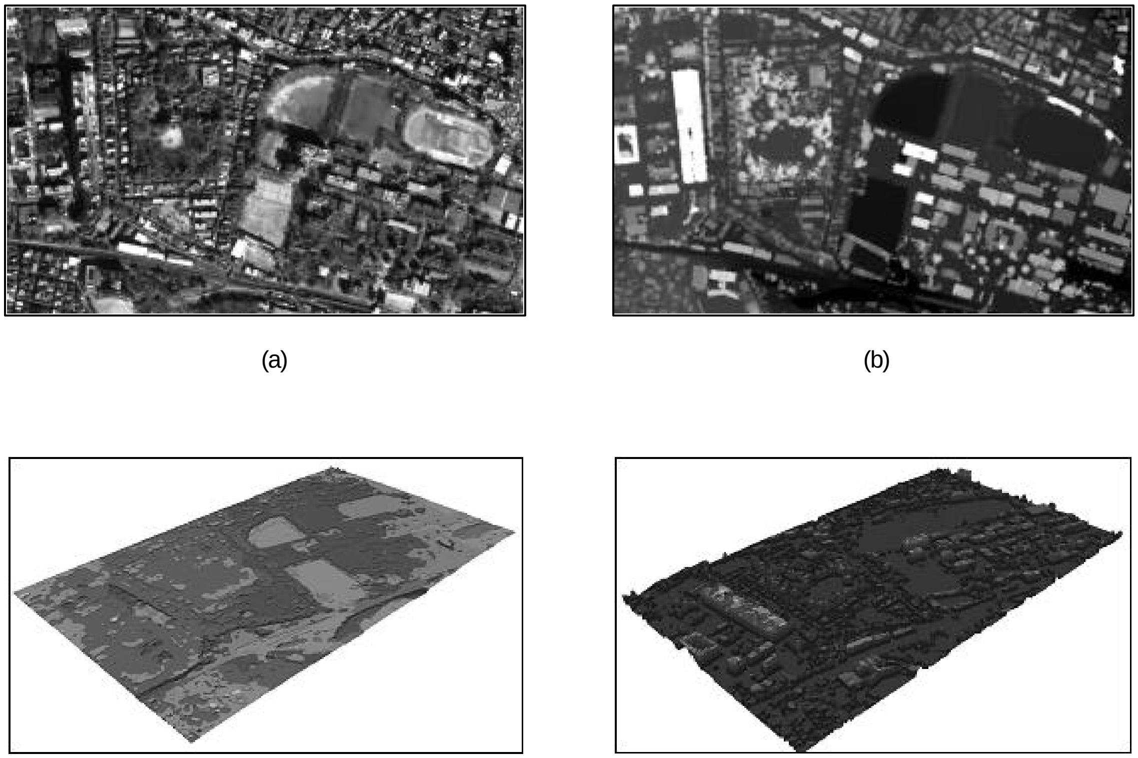
And it is necessary to verify the building contours, that is to separate building contours from those belong trees, lumps, etc. For the aerial images, the analysis of frequencies of edge direction as explained by Baltsavias et al (1995). may give the clue to separate regular shape from irregular shape. But for satellite images, the edge information is much blurred, detailed texture information will not be efficient in our case. As mentioned above, satellite images contains rich spectral information, it shows an opportunity to introduce spectral signature to help identify the object contained in detected contour, e.g. a statistic of NDVI will help to distinguish trees from small buildings. This method is also being used to split merged different contours and to connect separated but very closed contours in our system, because it is still under development, the results of this verification will not be presented in this paper.

3. EXPERIMENTS

Test site covers an area of 1000 x 600m over Komaba campus of University of Tokyo. This area contains most patterns of urban objects. Buildings in this area have various shapes and their sizes rang from very large to big to medium to small. The distribution of buildings ranges from sparse to very close. This area contains a lot of trees as well, from individual trees to tree crowds, some trees very close to buildings. Streets, roads, grasslands, sports fields and bare grounds are also contained in this test site.

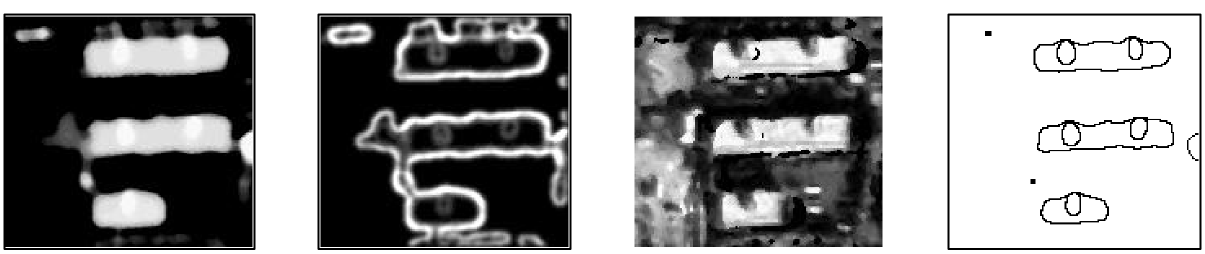
3.1 IKONOS images and ALS data

Our IKONOS images are kindly provided by Japan Space Imaging (JSI) Corporation. Acquisition date is on Nov. 4th, 2001. And images have been preprocessed by JSI to '*expert*' product level (http://www.spaceimaging.co.jp/seihin/seihin1.html), which is close to *Precision* products. All images have been rectified using ground control points (GCPs). The ground positioning accuracy (RSM) is 1.75m. The resolution of multi-spectral (R, G, B, NIR) is 4-meter and panchromatic image has 1-meter resolution. Due to the 11-bit color depth, IKONOS images contain rich spectral information. This will be useful for spectral detail analysis. In PAN image, most Objects can be recognized, but their boundaries are not easy to identify. All images are resampled into 0.5-meter GSD in order to co-process with DSMs. JSI also provides us ALS data, which originally captured using *ALTMS* system (manufacture is TerraPoint, built by HARC with NASA support). The average sampling space of laser points is about 1.5m and the size of footprint of laser beam on the ground is 90 cm, elevation accuracy is 15 cm and horizontal accuracy is about 1m. ALTMS system can record at most 4 multiple returns, but this information is not available in our ALS data. We simply generate a DSM from the ALS point clouds by Nearest Neighbor interpolation with 0.5 GSD the same as resampled IKONOS images.



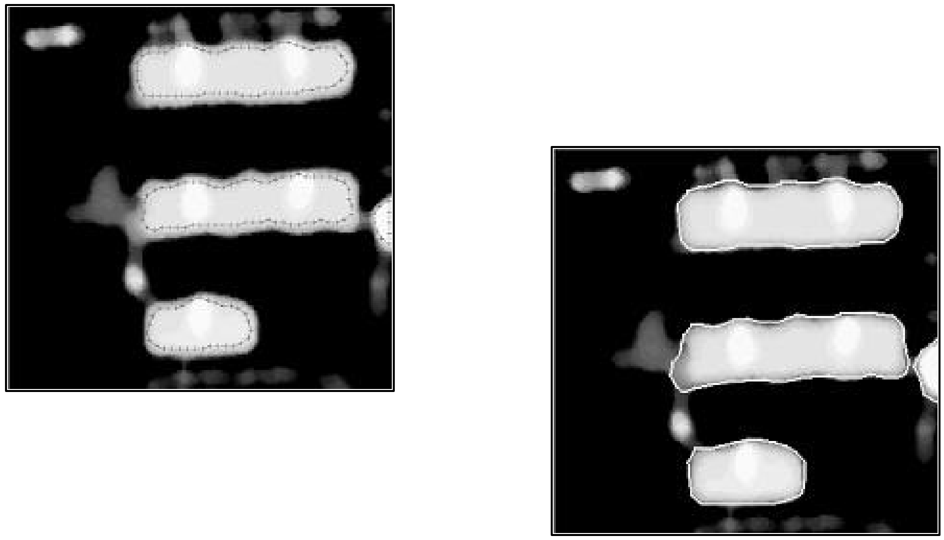
(a) (b) (c) (d)

Fig. 2: Test site data sets: (a) is IKONOS panchromatic band image (1 meter resolution); (b) is DSM generated from ALS data (0.5 meter GSD); (c) is 3D view of derived bare ground DEM; (d) is 3D view of derived normalized DSM.



(a) (b) (c)
(d)

Fig. 3: Balloon model for building contour detection: (a) nDSM; (b) DSM edges serve as one item of balloon inflation force (c) orthorectified IKONOS image and its gradients will contribute to balloon inflation force; (d) detected elevation blobs (12.0 –13.0 meter) using MHB method, it simplifies the processing and leads to relatively few target regions, and these blobs will serve as initial contours in balloon model.



(a) (b)

Fig. 7: Detected building contour using balloon model method: (a) elevation blobs as initial contours; (b) Initial contours move under constructed inflation force to approach boundaries, further verification needs to be conducted to extract 2D buildings from these potential boundary contours.

4. CONCLUSIONS AND FUTURE WORK

In this paper we present a snake-based approach for building extraction from urban high-resolution satellite images combining height data. The complementary characteristics of these two datasets are efficiently exploited from normalized DSM generation, objects segmentation and contours detection. This research shows that it is quite promising to combine spectral signature and height information for feature extraction. Using MHB method not only leads to a hierarchical representation of target objects, but also generates initial guesses for later automatic detection. Adapted balloon model takes clues from image and height as inflation force and the dynamic interpolation scheme makes contours to represent boundaries more accurately. However, although geometrical features didn't mentioned above, they are useful clues as well.

Future work will investigate the verification of detected contours by combining spectral signature with geometrical information, like straight lines, corners. A dynamic splitting and merging approach are under development to refine and verify contours in balloon model for automatic building extraction.

ACKNOWLEDGEMENTS

This research is carried out as a part of the joint research of Japan Space Imaging Corporation (JSI) and the Institute of Industrial Science (IIS) of the University of Tokyo. We would like to thank Mr. Akira Watanabe and Mr. Yunqing Li both from Japan Space Imaging Corporation, for kindly providing IKONOS imagery and ALS data used in this research.

REFERENCES

1. Baltsavias, E., Mason, S., Stallmann, D., 1995. Use of DTMs/DSMs and orthoimages to support building extraction. Swiss Federal Institute of Technology, Institute of Geodesy and Photogrammetry, http://e-collection.ethbib.ethz.ch/tagung/sg/02x.html.

2. Cohen, L, D., 1991. On active contour models and balloons. *Computer Vision, Graphics, and Image Processing: Image Understanding*, 53(2): P211-218, March 1991.

3. Fraser, C.S., Baltsavias, E., Gruen, A., 2001. 3D building reconstruction from high-resolution Ikonos stereo-imagery. In: *Automatic Extraction of Man-Made Objects From Aerial and Space Images (III)*, Edited by E.P.Baltsavias, A.Gruen, and L.V. Gool, Balkema, Tokyo, pp. 331-344.

4. Grun, A., 2000. Semi-automated approaches to site recording and modeling. *International Archives of Photogrammetry and Remote Sensing*, vol. 33, part B5, Amsterdam, pp. 309-318.

5. Haala, N., Brenner, C., 1999. Extraction of buildings and trees in urban environments. *ISPRS Journal of Photogrammetry & Remote Sensing*, 54, pp. 130-137.

6. Henricsson, O., 1996. Analysis of image structure using color attributes and similarity relations. Ph.D. Thesis, Swiss Federal Institute of Technology (ETH), Zurich, Diss.ETH, Nr.11663, Published in Mitteilungen, Nr.59 of the Institute of Geodesy and Photogrammetry.

7. Kass, M., Witkin, A., Terzopoulos, D., 1988. Snakes: Active contour models. *International Journal of Computer Vision*, (1): P321-331, 1988.

8. Lang, F., Forstner, W., 1996. 3D-city modeling with a digital one-eye stereo system. *International Archives of Photogrammetry and Remote Sensing*, vol. 31, part B3, pp. 415-420.

9. Mass, H.G., Vosselman, G., 1999. Two algorithms for extracting building models from raw laser altimetry data. *ISPRS Journal of Photogrammetry & Remote Sensing*, 54, pp. 153-163.

10. Murai, S., 2001. Generation of 3D city models in Japan – an overview. In: *Automatic Extraction of Man-Made Objects From Aerial and Space Images (III)*, Edited by E.P.Baltsavias, A.Gruen, and L.V. Gool, Balkema, Tokyo, pp. 59-64.

11. Tao, G., Yasuoka, Y., 2002. Combining high resolution satellite imagery and airborne laser scanning data for generating bareland DEM in urban areas. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol.XXX IV part No.:5/W3. In Cdrom.

12. Vosselman, G., Suveg, I., 2001. Map based building reconstruction from laser data and images. In: *Automatic Extraction of Man-Made Objects From Aerial and Space Images (III)*, Edited by E.P.Baltsavias, A.Gruen, and L.V. Gool, Balkema, Tokyo, pp. 231-239.

13. Weidner, U., Forstner, W.,1995. Towards automatic building extraction from high-resolution digital elevation models *ISPRS Journal of Photogrammetry & Remote Sensing*, 50(4), pp. 38-49.