An improved Snake for Automatic Building Extraction

LAU BEE THENG and RAYMOND CHIONG
School of Information Technology and Multimedia
Swinburne University of Technology Sarawak Campus
Jalan Simpang Tiga, Kuching, Sarawak
MALAYSIA

Abstract: - Using snakes to extract buildings from satellite imagery such as IKONOS and Quickbird is an active activity in remote sensing society. Building extraction from satellite imagery has more than 20 years of history but the automated extractions are still undergoing developmental stages due to increasing image variation, required level of details and higher resolution imagery are acquired. This paper discusses the prototyping of an improved snake model. Snake is initialized in cornered-radial cast in comparison to the existing centered-radial. Moreover the coefficients for snake energy are obtained through genetic algorithm.

Key-Words: - Snake contour, satellite imagery, genetic algorithm, building extraction.

1 Introduction

Generally, building extraction methods use generic models that assume all buildings follow a certain pattern. For this reason, generic models do not provide satisfactory extraction results when buildings in unstructured buildings like informal settlements precede the particular pattern (Ruther et al. 2002) in terms of building materials, neighborhood distance and orientations. Limited tools and methods are available to extract unstructured buildings compared to structured buildings.

2 Building extraction

Generally, the main tasks in building extraction from digital images are building detection and building reconstruction. However building extraction tasks may differ depending on the use of geometrical representation with rectangular models (Weidner and Frostner 1995), use of multiple images (Ballard and Zisserman 2000), and polyhedral shapes (Scholtze et al. 2002), the use of lines, points and regions to describe building outlines (Fisher et al. 1997). The existing automated building extraction techniques are still performing at elementary level caused by image variation in terms of type, scale, and required level of detail (Wang and Tseng 2003).

Recently some researches are published such as use of fused shadow data with 2D building blobs derived from normalized Digital Surface Model (DSM) (Li and Ruther 1999) and still video Kodak camera to extract shacks in South Africa (Baltsavias and Mason 1995). However DSM suffers from insufficient ground sampling data and matching errors due to poor image quality, and also occlusion and shadows that lead to poor definition of buildings outlines (Baltsavias and Mason 1995 and Ruther et al. 2002). An effective informal settlements extractor should accommodate both structured and unstructured buildings (Ruther et al. 2002). Automated recognition of object's semantic information is complicated as most algorithms fail whenever a new situation in image space is encountered such as objects are close to each other (Gruen 2000, Ruther et al. 2002). There are some researches on IKONOS and Quickbird satellite imageries. For instance, an automatic method of extracting buildings in densely urban areas from IKONOS images using large detached buildings without analysis of accuracy and structure details (Sohn and Dowman 2001); optimization and destruction approaches concurrently for building extraction where a point process technique is developed to extract well-structured buildings (Ortner 2002); comparisons of building extraction from IKONOS imagery with black and white aerial imagery to evaluate the potential of high-resolution images (Fraser et al. 2002); a research investigated the potential of Quickbird imagery for spatial data acquisition where sensors of 0.6m have lessened the gap between satellite images and aerial photographs that have resolution from 0.2 to 0.3m (Toutin and Cheng 2002); a linking edge chain is implemented to extract buildings from IKONOS images (Haverkamp 2003); high-resolution imagery for mapping urban areas in extracting land cover information from high-resolution images (Thomas et al 2003).

3 Snake

Snake or active contour is useful to extract structured and unstructured objects from informal settlements (Mayunga et al. 2005), tree crown extraction (Horvath et al. 2006), tracking dynamics (Sundaramoorthi et al. 2007), building extractions from IKONOS (Guo and Yasuoka 2003).

A snake, E_{Snakes} is represented by a vector, *V(s) = [x(s), y(s)]*, having the arc length *s* as a parameter, where *x* and *y* are the coordinates of a snakes contour point (Kaas et al 1988). The total energy of a snake is:

$$E_{Snakes} = \int_0^1 E_{Int}V(s) + E_{\mathrm{Im}g}V(s) + E_{Cont}V(s)ds \quad (1)$$

The solution of snake is activated by its intrinsic trend of minimizing its energies. The energy function reaches minimum when the snakes control points locks the object boundaries in the image space. There are several approaches implemented for building extraction related issues. For instance, the use of snakes and least squares method to extract buildings in 2D and 3D using aerial photography and satellite images (Mayunga et al. 2005). Cohen and Cohen (1990) used pressure force to control the movement of snakes contour. Although the method worked well, the parameter that controls the inflating force is difficult to estimate for high level noise in imagery. Tabb et al. (2000) combined snakes and neural networks to detect and categorize objects in images. The snake contour is stored as a vector of (x, y) coordinates reflecting the position of different control point on the contour's spline. The coordinates are the input for neural network.

Kreschner (2001) used homologous twin snakes and integrated in a bundle adjustment. The method fails when the system chooses wrong snakes contour. Ruther et al. (2002) use snakes and dynamic programming optimization technique to model buildings in informal settlement areas. However, dynamic programming is computational expensive and fails in more complex topologies. Guo and Yasuoka (2003) adopted a ballooning snake model. Multiple Height Bin (MHB) technique was employed to obtain the approximate snakes contour but it could not provide correct representation of the extracted objects.

4 Snake Initialization Problem

Mayunga et al. (2005) proposed an improved snake that discards weighted coefficients and external energy which creates boundary effects for unstructured buildings. The improved snake model (Equation 2) consists of only three energy terms that are Continuity, E_{Cont} (Equation 3), Curvature, E_{Curv} (Equation 4) and Image, E_{Img} (Equation 5).

The algorithm being minimized is expressed where $v_i = (x_i, y_i)$.

$$E_{Snakes} = E_{Cont} + E_{Curv} + E_{\mathrm{Im}g} \quad (2)$$

$$E_{Cont} = \sum_{i=1}^{n} \frac{|v_{i+1} - v_i|}{n} - |v_i - v_{i-1}| \quad (3)$$

$$E_{Curv} = |v_{i-1} - 2v_i + v_{i+1}|^2 \quad (4)$$

$$E_{\mathrm{Im}g} = 1 - \left| \frac{(g_{\max} - g_{v_i})}{(g_{\max} - g_{\min})} \right| |\cos\delta(v_i)|$$

where $g_{\max}$ and $g_{\min}$ denote the minimum and maximum gradient magniture of the contour point whereby g_{v_i} denotes the gradient magnitude (5)

Image term describes the radiometric content of the image and it restricts the snake points to move towards the points of highest gradient. The gradient of image at each control point is normalized to show small differences in values at the neighborhood of that control point. In this case, the gradient magnitude is negative to enable control points with large gradient to have small values. The image term attract the snakes to the image points with minimum gradient magnitude. Continuity term creates equal space snakes control points to avoid grouping and minimize distances. Curvature term expresses the curvature of the snakes contour.

To realize the model, Mayunga et al (2005) used a radial casting algorithm for initiating the snake model as shown in Fig.1. However, manual picking up of a building centre is essential before a radial cast can initialize a snake. Consequently, from the contour's centre, *C* radial lines are projected outwards at definable angular intervals as shown in Fig.1. Each snake control point in image space, advance to a new position where the gradient energy in a search window is maximum. The advancement has to proceed using angular intervals as well. The centre of the building polygon is always fixed and the radial distances, *l* to the snakes control points is variable depending on the size of the building object.

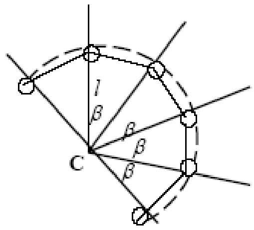

Fig.1 Radial casting

The problem with this radial case is the number of radial lines depends on the complexity of the building. Hence more complex structure would require a high volume of lines to form the snake contour. Furthermore the radial cast model claimed to be effective where a maximum of 16 radical lines should be sufficient to extract a building. However this is not feasible where unstructured building has irregular shape as shown in Fig.2. The extracted building, *C* varied from the original building, *A*, object *B* shows the 16 radial lines derived for the building from its centre.

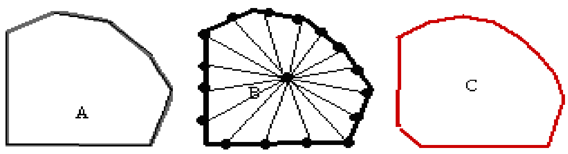

Fig.2 Radial casting of irregular shape

On the other hand, the snakes control points derived from centre possibly causes the snakes curve to be smaller than desired during radiation. If it happened, the generated snakes control points are deleted and a new centre is established. This cause inconsistency for all other snakes control points built unless they are all deleted and rebuilt.

5 Proposed Snakes Initialization

In response to the problems in existing radial casting, a proposed circular cast initialization for snake contour with genetic algorithm optimization is developed. A genetic algorithm is used to optimize the circular casting of irrelevant edges from satellite images. The building extraction problem is modeled as a finite-length binary string imitating a chromosome. Reproduction is a selection process in which individual strings containing the projected building boundary edge and the actual edge are kept to produce the next generation in accordance with their fitness values. The finalized building boundary edges are generated from the parents based on a randomly generated crossover process. Let's look at how circular casting work followed by GA optimization.

5.1 Corner point initiation

The first control point, *C* can be initiated at any corner point of the building object not necessarily the center as picking up center requires manual operator. Hence this allows automation where the point can be picked by comparing control points using corner detector. Corner detector works by searching for pixel in an object based on the assumption that corners are associated with maxima of the local autocorrelation function. It calculates corner value for each pixel of the image, and if *C* is a local maximum above predefined threshold, the pixel is declared the first control point, *C*. Consequently a cast is initiated from corner as control point as shown in Fig.3. The circular cast resolves the inability of radial cast to cope for irregular shape as proposed in Mayunga et al. (2005) as shown in Fig.2. In fact, 8 radial lines derived for the building from its first control point is sufficient to determine the irregular shape in Fig.3.

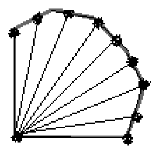

Fig.3 Corner point initiation and casting for irregular building shape

```
------------------------------------------------------------
Begin
Generation = 0
Set control point = local maxima
while extraction is not closed
   create circular cast
   while there is more pixel  within circular cast
        get pixel
        if pixel attribute = local maxima
            append pixel as a node to snake contour
        end if
 end while
end while
------------------------------------------------------------
```

Fig.4 GA optimized circular cast

5.2 GA optimization

As there is an existence of several weighted coefficients for the calculation of snake energy, they are attempted in a trial-error manner. GA is an evolutionary method of optimization and it produces good results for this kind of parametric optimization. Now, the most important aspect concerning circular casting of buildings in snake model using GA is to select the appropriate population and fitness function. Selecting proper population reduces searching time in the optimization process and selecting proper fitness function lessens the convergence time. Consequently, the searching region, *r* to detect the contour of object is decided with GA.

5.3 Fitness Functions

The fitness function uses the least-squares solution for retrieving optimal solution based on an assumption that contour pixels screened in the circular cast correspond with the real contour of the object. Consequently, the number of corresponding control points captured by the circular cast is a crucial factor. Hence a highly structured fitness function is used in Equation 6 where the fitness will be a maximum if energies are minimal. γ, α, and β are the coefficients of fitness. Equation 7 shows the string architecture of the chromosome that produces a string of real numbers. n is the number of control point, d_i is the distance from corner of object to control point, α_i is the angle crossing point and horizontal axis, m is the number of population. Each extracted building shape is treated as s chromosome. If a chromosome represents a set of contour points varied from the ground truth object, it may lead to either too many irrelevant or too few relevant contour points are captured correctly. It can either make a correct indication or mislead the extraction.

$$Fitness = 1 + \frac{1}{\gamma E_{img} + \alpha E_{cont} + \beta E_{curv}} \quad (6)$$

$$\begin{aligned} s_1 &= \{d_1, d_2, \ldots, d_n, \alpha_1, \alpha_2, \ldots, \alpha_n\} \\ s_2 &= \{d_1, d_2, \ldots, d_n, \alpha_1, \alpha_2, \ldots, \alpha_n\} \quad (7) \\ s_m &= \{d_1, d_2, \ldots, d_n, \alpha_1, \alpha_2, \ldots, \alpha_n\} \end{aligned}$$

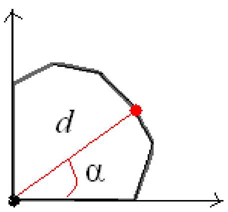

Fig.5 Distance from corner of object to control point, *d* and angle crossing control point and horizontal axis, *α* for building shape

6 Testing

The GA optimized snake initialization for buildings extraction from satellite imagery is prototyped. The prototype consists of image pre-processing and building extraction. In the image preprocessing stage, variation in illumination conditions, shadowing and building density in urban areas makes it very difficult to distinguish individual buildings from its surrounding. In order to solve this problem, a non-linear anisotropic diffusion model (Weickert 1999) was used to normalize the noise effects around the buildings. The image normalization process brings the variation of points around the buildings at the same level. The diffused image is then used as an input for building extraction.

Once a corner point is obtained for a building in the image space and then the snakes contour are automatically generated. User is given an option to accept or reject a single snake contour or all generated snake contours. If a snake contour is accepted, an iterative minimization function is invoked, minimum and maximum energy values in the neighborhood are computed. Neighborhood point with the lowest energy value is the new position in the image space. The iteration stops when snake contour completes/locks a building outline. A testing was done on satellite imagery for informal settlements at Burlington city in Fig.6. The image used was obtained from the website and geo-rectified by the image vendor DigitalGlobe, it has 4-band spectral resolution at 16 bits/pixel, pre-processed by vendor and geometrically corrected.

7 Quantitative Analysis

A total of 100 building corner points from 2D vectorized layer were randomly compared with their corresponding points from the ground truth data. Fig.7 shows extracted buildings in informal settlements.

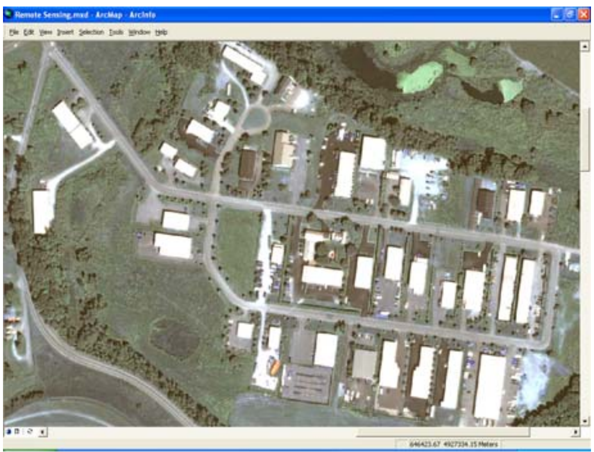

Fig.6 A Quickbird imagery of Burlington area

Fig.7 Extracted buildings from the test area

Table 1 shows the number of population, crossover rate, number of generation, mutation rate, and coefficient of fitness used in the testing. This testing has a population of 100 buildings, crossover rate of 0.60 and mutation rate of 0.01 taken from generic practice in satellite imagery processing (Goldberg 1989). The results of fitness across generation in Fig.8 show that the snake contour initiation has been optimized and the extracted buildings are closer to the ground truth objects. On the other hand, Root Mean Square Error (RMSE) was computed to determine the internal accuracy of the measurement. Standard deviations in (x, y) were also computed in Table 2. Model 1 represents the results obtained from this research whereby Model 2 refers to the traditional improved snake model with radial casting (Mayunga et al. 2005). The possible reason contributing to this figure is closeness of buildings in informal settlements and resolution of the image. Informal settlements consists of random noise causes edges along the corners to divert from their correct positions. There is a need for post-processing stage to refine the edges if higher accuracy application is required.

Table 1 Coefficients of genetic algorithm

No. of population, *m*	100
Crossover rate	0.60
Mutation rate	0.01
No. of generation	150
Coefficient of fitness:	
[1] γ	1.2
α	0.1
β	0.1

8 Conclusion

The proposed GA optimized circular casting snake initialization improved the RMS and standard deviations from the ground truth data. It is a significant contribution to building extraction from high-resolution satellite imagery. The approach has been tested on informal settlements where buildings with irregular shapes are extracted with reliable accuracy.

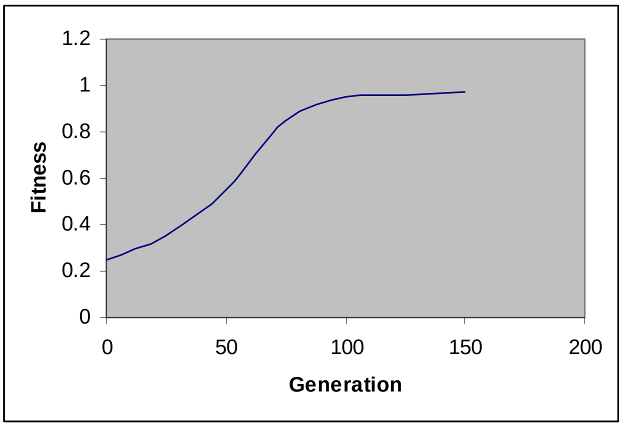

Fig.8 Fitness across generation

Table 2 RMSE, deviations of contour points

Model	No. of points	RMSE (m)	Std dev. in x	Std dev. in y
1	20	0.9	0.5	0.8
2	20	1.22	0.80	0.98

References

[2] Ballard, C. and Zisserman, A. (2000). A Plane-sweep Strategy for the 3D Reconstruction of Buildings from Multiple Images. *IASPRS*, Vol. XXXIII, pp 56-62.

[3] Baltsavias, E. and Mason, S. (1995). Image–Based Reconstruction of Informal Settlements. Automatic Extraction of Man Made Objects from Aerial and Space Images II, Birkhauser Verlag, Berlin, pp. 97-108.

[4] Cohen, L. D. and Cohen, I. (1990). A Finite Element Method Applied to New Active Contour Models and 3-D Reconstruction from Crossing Sections. In *Proceedings of the Third International Conference on Computer Vision*, Osaka, pp. 587-591.

[5] Digital Globe Inc. QuickBird Sample Imagery. Accessed 30/03/06. www.digitalglobe.com/sample_imagery.shtml.

[6] Fraser C. S., Baltsavias, E. and Gruen A. (2002). Processing of IKONOS Imagery for Submeter 3D Positioning and Building Extraction. *International Journal of Photogrammetry and Remote Sensing* Vol.

56 pp. 177-194.

[7] Fisher, A., Kolbe, T. H. and Lang, F. (1997). Integration of 2D and 3D Reasoning for Building Reconstruction Using a Generic Hierarchical Model, In *Proceedings of the Workshop on Semantic Modeling for the Acquisition of Topographic Information from Images and Maps SMATI'97*, pp. 159-180.

[8] Goldberg, D. E. (1989). *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley, USA.

[9] Guo, T. and Yasuoka, Y. (2003). Snake-based Approach for Building Extraction from High-Resolution Satellite Images and Height Data in Urban Areas. Accessed 30/03/06. http://www.gisdevelopment.net/aars/acrs/2002/vhr/018.pdf

[10] Gruen, A. (2000). Potential and Limitation of High-Resolution Satellite Imagery. *Proceedings of the Asian Conference on Remote Sensing,* December 4-8, Taiwan.

[11] Haris, C. and Stephens, M. (1988). A Combined Corner and Edge Detector. In *Proceedings of 4th Alvey Vision Conference*, Manchester, pp. 141-151.

[12] Haverkamp, D. (2003). Automatic Building Extraction from IKONOS Imagery, http://www.spaceimaging.com/whitepapers_pdfs /2004/

[13] Kaas, M., Witkin, A. and. Terzopoulos, D. (1988). Snakes: Active Contour Models. *International Journal of Computer Vision*, (1): p321-331.

[14] Kreschner, .M. (2001). Homologous Twin Snakes Integrated in a Bundle Block Adjustment. Accessed 30/03/06. http://www.ipf.tuwien.ac.at/publications/mk_p_isprs98/paper3.html

[15] Li, J. and Ruther, S. H. (1999). IS-Modeller: A low-Cost Image-based Tool for Informal Settlements Planning, *Geoinfomatics '99 Conference*, Ann Arbor, 19-21 June, 1999, pp. 1-9.

[16] Mayunga, S. D., Zhang, Y. and Coleman, D. J. (2005). Semi-Automatic Building Extraction Utilizing Quickbird Imagery. In *Stilla, U., Rottensteiner, F. and Hinz, S. (Eds.), IAPRS*, Vol. XXXVI, pp. 131-136.

[17] Otner, M., Descombes, X. and Zerubia J. (2002). Building Extraction from Digital Elevation Model, *Rapport INRIA No. 4517*, INRIA France.

[18] Perez, M. M. and Dennis, T. J. (1997). Adaptive Implementation of the SUSAN Method for Image Edge and Feature Detection. In *Proceedings of IEEE Conference on Image Processing*, pp. 394-397.

[19] Ruther, H., Hagai, M. M and Mtalo, E.G. (2002). Application of Snakes and Dynamic Programming Optimization in Modeling of Buildings in Informal Settlement Areas, *ISPRS Journal of Photogrammetry and Remote Sensing*, 56(2002): p 269-282.

[20] Scholze S., T., Moons, T. and Gool, V. (2002). A Probabilistic Approach to Roof Extraction and Reconstruction. *IASPRS*, Commission III, pp. 231-237.

[21] Sohn, G. and Dowman, I. (2001). Extraction of Buildings from High Resolution Satellite Data. In *Baltsavias, E., Gruen, A., Van Gool, L. (Eds.), Automated Extraction of Man-Made Objects from Aerial and Space Images (III)*. Balkema Publishers, Lisse, pp. 345-355.

[22] Tabb, K., Davey, N. and Stella, G. (2000). Analysis of Human Motion using Snakes and Neural Networks. Accessed 30/03/06. http://homepages.feis.herts.ac.uk/~nngroup/pubs/papers/tabb-amdo00.pdf

[23] Thomas, N., Hendrix, C. and Cogalton, R.G. (2003). A Comparison of Urban Mapping Methods Using High-Resolution Digital Imagery. *Photogrammetric Engineering and Remote Sensing,* Vol. 69, No. 9, September 2003, pp. 963-972.

[24] Toutin, T. and Cheng, P. (2002). 3D Models for High Resolution Images: Examples with QuickBird, IKONOS and EROS. *Symposium on Geospatial Theory, Processing and Applications*, Ottawa.

[25] Tseng, Y., Lin, C. and Wang, S. (2002). Model-image Fitting Using Genetic Algorithm for Building Extraction from Aerial Images. In the *proceedings of ACRS2002*.

[26] Weidner, U. and Forstner, W. (1995). Towards Automatic Building Extraction from High Resolution Digital Elevation Models. *International Society of Photogrammetry and Remote Sensing,* Vol. 50 No. 4, pp 38-49.

[27] Wang, S. and Tseng, Y.H. (2003). Semi-automated Building Extraction Based on CSG Model-Image Fitting. *Photogrammetry Engineering & Remote Sensing*, 69:2, pp. 171-180

[28] Weickert, J. (1999). Coherence-enhancing Diffusion Filtering. *International Journal of Computer Vision*, Vol. 31 pp. 111-127.