Semi-automatic Building Extraction from Quickbird Imagery

Selassie David Mayunga
School of Geospatial Sciences and Technology, Ardhi University, Dar Es Salaam 35176, Tanzania

Abstract: Automatic extraction of features and buildings in particular from digital images is one of the most complex and challenging tasks faced by computer vision and photogrammetric communities. Extracted buildings are required for varieties of applications including urban planning, creation of GIS databases and development of urban city models for taxation. For decades, extraction of features has been done by photogrammetric methods using stereo plotters and digital work stations. Photogrammetric methods are tedious, manually operated and require well-trained personnel. In recent years, there has been emergence of high-resolution space borne images, which have disclosed a large number of new opportunities for medium and large-scale topographic mapping. In this paper, a semi-automatic method is introduced to extract buildings in planned and informal settlements in urban areas from high resolution imagery. The proposed method uses modified snakes model and radial casting algorithm to initialize snakes contours and refinement of building outlines. The extraction rate is 91 percent as demonstrated by examples over selected test areas. The potential, limitations and future work is discussed.

Key words: High-resolution imagery, building extraction, informal settlements, snakes models.

1. Introduction

Automatic extraction of features from digital images is one of the most complex and challenging task faced by computer vision and photogrammetric communities. Feature extraction and buildings in particular are required for varieties of applications such as urban planning, creation and updating of Geographic Information Systems databases and creation of urban city models for taxation. In practice, the extraction of buildings from digital images is complex because buildings particularly in dense urban areas, which have complex forms and roofs of various compositional materials. For decades’ extraction of features in the developing world, it has been manually used stereo plotters or occasionally digital workstations. Manual based building extraction is slow, expensive and requires well-trained operators. However, for rapid urbanizing and high-densely urban areas, it becomes even more difficult [1].

Numerous efforts have been made in the past two decades to automate building extraction from digital images [2-4]. Fully automatic feature extraction systems are limited to specific applications and not yet operational [5]. In the absence of fully automatic systems, semi-automatic systems seem to be an alternative solution [6] for feature extraction. In recent years, there has been emergence of high-resolution space borne images, which have shown potential for medium and large-scale mapping in urban areas [7]. In order to effectively utilize the availability of high resolution satellite images, new techniques and tools are urgently required.

In this paper, an effective semi-automatic method is introduced to extracts buildings in planned and informal settlements in urban areas by measuring a single point on the approximate centre of the building. Then, after a radial casting, algorithm is invoked to initialize the snakes contour and refinement of the building outlines.

1.1 Related Works/Previous Work

In an effort to make building extraction processes efficient, various attempts to automate building
Semi-automatic Building Extraction from Quickbird Imagery

extraction processes have been reported in the past two decades. However, existing automated building extraction techniques are still operating at a very rudimentary level [8]. One first difficulty for automated building extraction is caused by image variation in terms of type, scale and required level of detail [9]. Secondly, automatic recognition of semantic information of an object using computers is very difficult as existing automatic algorithms tend to fail whenever a new situation on the image is encountered. A well-extracted building requires well interpretation of the image to recognize its location and extent, and then automated processes are employed.

Sohn, G. and Dowman, I. [10] proposed an automatic method of extracting buildings in densely urban areas from IKONOS images. In their study, large detached buildings were used but there was no analysis on the accuracy and modeling of the structures. Fraser, C. S. et al. [11] compared buildings extracted from IKONOS imagery with those obtained using black and white aerial photographs to evaluate the potential of high-resolution images. Toutin, T., and Cheng, P. [12] investigated the potential of quickbird image for spatial data acquisition and showed that quickbird sensors of 0.6 m have narrowed the gap between satellite images and aerial photographs for large scale mapping. In the same perspective, Thomas, N., Hendrix, C., and Goglon, R. [13] assessed three different classification methods for extracting land cover information from high-resolution images. In their investigations, it was concluded that high-resolution satellite imagery is a valuable tool for large scale mapping urban areas. In the absence of automatic building extraction systems, semi-automatic systems seem to be an alternative solution for feature extraction [14].

Semi-automatic building extraction methods have been investigated by photogrammetry communities and computer vision experts for more than two decades, although most of the existing methods are limited to specific applications. Brunn, A., and Weidner, U. [15] used building detection and building reconstruction techniques to extract buildings, however, these tasks are tedious as cannot be combined. The application of geometrical representation with rectangular models developed by Weidner, U., and Ballard, C. et al. [16, 17] used multiple images and polyhedral shapes to describe building outlines. Generally, most of the existing semi-automatic methods referred above work well where buildings are assumed to follow a certain pattern or code. Therefore, using such methods in areas where buildings does not follow any particular pattern or code especially in informal settlements areas cannot provide realistic results [18]. Buildings in informal settlements areas are made of diverse materials, very close to each other, have complex structures and have no proper orientation which makes the extraction process very difficult. Informal settlements are commonly found in the developing world and accounts from 60 percent to 70 percent of buildings in urban areas. In Tanzania, for example, about 70 percent of buildings in urban areas are in informal settlements.

There have been very limited attempts to develop tools and methods to extract buildings in informal settlements areas as compared to the research efforts made to extract buildings in planned settlements. Recent efforts include collection of social and spatial information, used fused shadow data with 2D building blobs derived from normalized Digital Surface Model (DSM) and the use of still video camera (DCS460c) to extract shacks in South Africa [19]. A method developed by Li, J., and Ruther, S. H. [20] used DSM, shadow and linear feature data derived from low-cost small-format digital imagery to extract buildings in informal settlements.

A common feature for building extraction methods referred above, achieves their extraction process utilizes a DSM or DEM. The main disadvantages of using DSM and DEM generated by image matching technique includes insufficient ground sampling data
and matching errors caused by poor image quality, occlusion and shadows which leads to poor definition of buildings outlines [21]. For the technology to be effective, it must be able to extract buildings both in planned and informal settlements. Active contour models or commonly known as “snake’s models” at present seem to be an alternative and useful model to extract buildings in informal settlements.

1.2 Existing Snake’s Models

Active contour models also known as “snake’s model” was first introduced by Kaas, M., Witkin, A., and Terzopoulos, D. [22]. In the theory of snakes, the control points on the closed curve are guided by active contour models, which minimized a weighted combination of internal, image and external energy forces. The internal energy force describes the shape of the snake. The image force attracts the snakes to the boundaries of the object, while the external energy force comes from the image itself or a higher level of image processing. The solution of the snake’s models are activated by its intrinsic trend of minimizing its energies. The function is defined in such a way that its energy reaches minimum when the snakes control points locks the boundaries of an object.

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

$$E_{ext} = E_{img} + E_{con}$$

Where, $E_{img}$ is the image energy, which can be the image intensity or intensity gradient, and $E_{con}$ is external constraint, variable constraints which can be introduced into snake’s model. For each control point on snake contour, its total energy can be represented as described in Eq. (4).

$$E = \alpha E_{conti} + \beta E_{curv} + \pi E_{int} + \mu E_{cont}$$

Where, $\alpha, \beta, \pi$ and $\mu$ are the weights of every kind of energy.

1.3 Limitation of Existing Snake’s Models

There are limitations on the use of snake’s model for building extraction which have not yet completely solved. For example, it is difficult to determine the appropriate weighted coefficients of the energy functions which cause bunching up of snake’s points on an image. Also, there is no simple way of initializing snakes contours. Several approaches have been proposed to remedy the above-mentioned limitations. For example, Trinder, J., and Li, H. [23] used snake’s model and least squares to extract buildings in 2D and 3D using aerial photography and satellite images. Cohen, L. D. [24] used pressure force to control the movement of snake’s contour. Although these modified methods works well in many cases, but the parameters that controls the inflating force of the contour is not easy to set especially for high level of noise in the image.

Tabb, K. [25] combined snakes and neural networks to a different position of control point on the contours detect and categorize objects in images. In this approach, a snake’s contour is stored as a vector of (x, y) coordinates and each (x, y) coordinate reflecting spline. Then the coordinates are used as input into the neural network.

Kreschner, M. [26] used homologous twin snakes and integrated in a bundle adjustment to extract buildings. This technique often fails when a wrong
snake’s contour is selected. Ruther, H., Hagai, M., and Mtalo, E. G. [27] used snakes and dynamic programming optimization technique to model buildings in informal settlement areas. However, the dynamic programming is computational expensive and occasionally fails to determine the exact shape of the buildings. Guo, T., and Yasuoka, Y. [28] adopted “balloon snake’s model” with Multiple Height Bin (MHB) technique to obtain the approximate snakes contour. However, the MHB technique could not provide correct representation of the extracted building contour.

1.4 Improved Snake’s Model for Energy Minimization Function

In attempt to solve the limitations of snake’s model as described above, an improved snake’s model is introduced whereby a snakes external energy term that is disregarded, which creates boundary effects for buildings. Instead, the weighted coefficients $\alpha$, $\beta$, $\pi$ and $\mu$ are fixed to allow a uniform strength of the snakes energy terms. The improved snake’s algorithm for building extraction in planned and informal settlements is represented in Eq. (4):

$$E = E_{cont} + E_{img}$$

Where, $E_{cont}$ and $E_{curv}$ are energy terms as expressed in Eq. (1). The first and second internal energy terms are briefly described in discrete form:

1. Continuity term

Let $v_i = (x_i, y_i)$ be a snake’s control point on an image space, from Eq. (4), continuity term is expressed as in Eq. (5):

$$E_{cont} = \bar{t}|v_i - v_{i-1}|$$

Where, $\bar{t}$ is the mean distance between two snake’s control points and it is expressed as in Eq. (6):

$$\bar{t} = \frac{\sum_{i=1}^{n} v_i - v_{i-1}}{n}$$

Where, $n$ is the number of control points. This term constrained the snake’s control points to have equally spaced avoiding points to be grouped in one point and at the same time, minimizing the distance among these points.

2. Curvature term

This term expresses the curvature of the snake’s control points and smoothness of the snake’s contour and mathematically is defined as in Eq. (7):

$$E_{cur} = |v_{i-1} - 2v_i + v_{i+1}|$$

(7) Image term

The image term describes the radiometric content of the image and it restricts the snake’s control 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. An expression of the image term is defined as in Eq. (8):

$$E_{img} = \frac{\text{Min}-\text{Mag}}{\text{Max}-\text{Min}}$$

(8) Where, $E_{img}$ is the image energy term, $\text{Min}$ is the minimum gradient magnitude in the neighborhood $\text{Mag}$ is the gradient magnitude at each control point, and $\text{Max}$ is the maximum gradient magnitude in the neighborhood. The image energy terms described above attract the snakes to the image points with minimum gradient magnitude.

1.5 Radial Casting Algorithm

In order to overcome the limitations of snake’s model, a radial casting algorithm is developed to initialize the snake’s control points. A single seed point called Centre (C) is measured at approximate centre of each building. The radial lines from this point grow radically to lock the building outlines. The radial casting lines are shown in Fig. 1.

For each seed point (C):

1. The contour’s centre point C is measured and from this point, radial lines are projected outwards at definable angular intervals. The angular interval consists of four, eight or sixteen radial lines ranging from $0^\circ$-360. The number of radial lines depends on the complexity of the building;
(2) The distances of radial lines joining the central point C in the snake’s contour is computed. The radial line is termed as the I’s line;

(3) The centre point C of the building polygon is always fixed and the radial distances to the snake’s control points which is variable depending on the size of the building object. Several different radial lines were tested to identify an optimal number of lines suitable for building extraction in informal settlement areas. Following this experiment and based on the dynamics of buildings in informal settlements, 8 and 13 radial lines for simple and complex buildings respectively were adopted as an optimal number of lines for this application;

(4) Each snake’s control point in image space, advance to a new position where the gradient energy in a search window is maximum.

During radiation process of the snake’s control points from point C, it is possible that the snake’s curve becomes smaller than desired. If this happens, the generated snake’s control points can be deleted and a new centre point C can be established.

2. Materials

In this method, there are two main processes: image pre-processing and building extraction.

2.1 Image Pre-processing

To process high-resolution satellite images for subsequent building extraction, the operation aspects of the image acquisition have to be taken into consideration [29]. These aspects have effects on the homogeneity or non-homogeneity of image quality particularly in high-densely built-up areas. The image quality is mostly affected by variations in sensor view angle, sun angle, shadowing and atmospheric conditions [30]. These effects become worse in areas, where buildings roofs have various composition materials. For example, high-resolution images with 8 bits have a loss of information in shadow or in bright areas [31]. Ross, L. [32] discussed difficulties commonly observed when dealing with urban shadowed areas. In his study, he recommended the use of 11-bits image in order to improve visual interpretation of objects.
The variation in illumination conditions of the image, shadows and building density in urban areas makes it very difficult to distinguish individual buildings from its surrounding particularly in informal settlements. In order to solve this problem, a non-linear anisotropic diffusion model was adopted and implemented to normalize the noise effects around the buildings [33]. The diffusion process establishes a scale-space which normalizes image noise concentration. The aim of image normalization is to bring the variation of pixels around the buildings at the same level. The diffused image is then used as an input into the modified snake’s algorithm for subsequent extraction of the building outlines.

2.2 Building Extraction

To effectively extract buildings using snake’s models and radial casting algorithm, an operator measures a single point at the approximate at the centre of the building in the image space and then the snake’s points along the contour are automatically generated. As soon as the snake’s points are generated, the operator has an option to accept or reject the snake’s contour. In the event of rejecting a snake’s contour, a single snake’s contour or all generated snake’s contours can be deleted. Conversely, if the snake contour is accepted, then a minimization function is invoked, and for each snakes, control point at $3 \times 3$ search window whereby the minimum and maximum energy values in the neighborhood are computed. This process is iterative and at the end, the point with a minimum energy is selected as a new position in the image space and a final solution is reached when snake’s contour locks the building outline.

3. Results and Discussion

3.1 Data Used and Studied Area

The method developed in this study was applied to extract buildings from informal settlements in Dar Es Salaam city, Tanzania and on planned area of Oromocto, Township in New Brunswick to compare the effectiveness of this method.

3.2 Building Extraction Results

The snake’s model and radial casting algorithm have been implemented whereby Fig. 2a shows extracted buildings in informal settlement and Fig. 2b shows the result of building extraction in vector layer. Fig. 3 shows extracted buildings in planned area.

It is worth to mention that most of the buildings outline from informal settlements (Fig. 2) and planned settlements (Fig. 3) areas were well extracted.

3.3 Analysis of the Results

In this study, the qualitative and quantitative analysis was performed to compare the extracted buildings using snakes and radial casting algorithm with the results obtained by manual plotted using photogrammetric analytical plotter.

3.3.1 Qualitative Analysis

In the qualitative analysis, the objective was to determine the practicability of the proposed approach whereby a building extracted percentage rate which is calculated. For this metric, a modified form of approach used by Avrahami, Y. [34] has been applied. In the model, a weighted parameter $k = 0.5$ for buildings partially mapped was used to compute the extraction percentage rate. Since the extraction process was carried out in the same environment, each parameter in the model has equal contribution to the final computation of the building extraction rate. In the modified form, the building extraction rate is expressed as in Eq. (9):

$$BER = \frac{BCE}{BCE+BPE+BNE}$$

Where, $BER$ is the Building Extraction Rate, $BCE$ is the Building Correctly Extracted, $BPE$ is the number of Buildings Partially Extracted and $BNE$ is number of Buildings not extracted. Table 1 presents the extraction rate in each test area. These results show that from three test areas, buildings were extracted at 93.6 percentage rates.
3.3.2 Quantitative Analysis

In quantitative analysis, building corner points from 2D vector layer were randomly selected and measured. The measured points were compared with their corresponding points measured from photogrammetric method. A total of 20 points were measured from test area 1 as shown in Table 2.

From 20 randomly measured points, the mean Root Mean Square Error (RMSE) was computed to determine the internal accuracy of the measurement. The standard deviations in x and y were also computed as summarized in Table 2. The results showed that the standard deviations are 0.80 m and 0.98 m in x and y respectively. However, it is important to mention that the proposed method has not been able to clearly define building corners for some buildings in the same

Fig. 2  (a) Extracted buildings from informal settlements and (b) Extracted buildings polygons (from vector layer).
Fig. 3 Extracted buildings from planned settlements (Oromoncto area).

Table 1 Building extraction percentage rate from three test areas.

<table>
<thead>
<tr>
<th>Test area</th>
<th>BCE</th>
<th>BPE</th>
<th>BNE</th>
<th>Extraction rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>74</td>
<td>4</td>
<td>2</td>
<td>92.5</td>
</tr>
<tr>
<td>2</td>
<td>36</td>
<td>1</td>
<td>1</td>
<td>94.7</td>
</tr>
</tbody>
</table>

Table 2 The RMSE and deviations of randomly measured building corner points.

<table>
<thead>
<tr>
<th>Test areas</th>
<th>No. of points</th>
<th>RMSE (m)</th>
<th>Std. Dev. in (x)</th>
<th>Std. Dev. in (y)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dar es salaam test area</td>
<td>20</td>
<td>1.22</td>
<td>0.8</td>
<td>0.98</td>
</tr>
</tbody>
</table>

Table 3 Time used to extract a single building in the test areas.

<table>
<thead>
<tr>
<th>Time used to extract a single building</th>
<th>Time used for each building (in seconds)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Semi-automatic</td>
</tr>
<tr>
<td>Scene navigation</td>
<td>34</td>
</tr>
<tr>
<td>Building extraction</td>
<td>10</td>
</tr>
<tr>
<td>Total time used</td>
<td>44</td>
</tr>
</tbody>
</table>

manner as appeared from ground truth data. The possible reason could be the closeness of buildings in informal settlements as well as the resolution of the image. The higher random noise effects in high-resolution imagery causes edges of the building along the corner of buildings to wobble from their correct positions.

3.3.3 Comparison on Time Used to Extract Building

The time used to extract a single building consists of the parameters:

(1) Scene navigation;

(2) Building extraction.

The aim of this test was to provide an indication of the efficiency of extracting a single building between modified snake’s model and manual system. Scene navigation time is recorded when the human operator is interacting with the image before building extraction process and building extraction time which is the actual time used to extract a single building. The recorded time during experiment is illustrated in Table 3.
Based on the same level of details extracted between manual and semi-automated processes, it was revealed that the semi-automated process developed in this study reduced the time to extract a single building by about 23%.

4. Conclusions

This paper demonstrates that the modified snake’s models and radial casting algorithm improved the extraction rate by 23%, and delivered a significant result of building extraction from high-resolution satellite imagery. This method shows that, for all tests, areas buildings with different shapes and orientation were extracted with reliable accuracy.

In addition, the minimum time used to extract a single building using conventional photogrammetric method is 57 seconds as reported by Ruzgienė, B. [35]. However, by using snake’s models and radial casting algorithm, a single building can be extracted for 44 seconds only. This extraction time is significantly smaller as compared to conventional photogrammetric technique. Therefore, it can be reported here that building extraction using modified snake’s model and radial casting algorithm can be practically used in planned as well as in informal settlements areas.

References

Semi-automatic Building Extraction from Quickbird Imagery


