Computers, Environment and Urban Systems 34 (2010) 435–441

Contents lists available at ScienceDirect

Computers, Environment and Urban Systems

journal homepage: www.elsevier.com/locate/compenvurbsys

An improved snake model for automatic extraction of buildings from urban aerial images and LiDAR data

Mostafa Kabolizade *, Hamid Ebadi, Salman Ahmadi

Department of Photogrammetry and Remote Sensing, Faculty of Geodesy and Geomatics, K.N. Toosi University of Technology, P.O. Box 1996715433, Tehran, Iran

ARTICLE INFO

Article history:
Received 31 January 2009
Received in revised form 29 April 2010
Accepted 30 April 2010

Keywords:
Building extraction
Snake model
LiDAR data

ABSTRACT

The automatic extraction of objects from data and images has been a topic of research for decades. This paper proposes an improved snake model that focuses on building extraction from color aerial images and light detection and ranging (LiDAR) data. A snake is defined as an energy minimizing spline guided by external constraint forces and influenced by image forces that pull it toward features such as lines or edges. Based on the radiometric and geometric behaviors of buildings, the snake model is modified in two areas: the criteria for the selection of initial seeds and the external energy function. The proposed snake model includes a new height similarity energy factor and regional similarity energy as well as gradient vector flow (GVF), which efficiently attracts the snake approaching the object contours. Compared with the traditional snake model, this algorithm can converge to the true building contours more quickly and more stably, especially in complex urban environments. Examination of the results shows that buildings extracted from a dense and complex suburban area using the GVF model have an 81% shape accuracy, whereas the improved model has a 96% shape accuracy.

1. Introduction

The automatic extraction of man-made objects such as buildings and roads from imagery taken in urban areas has become a topic of growing interest for photogrammetry, computer vision and remote sensing. Research in this area started in the late 1980s; many different types of source images ranging from single intensity images, color images and laser range images to stereo and multiple images were used (Peng, Zhang, & Liu, 2005). Some useful applications of this technique involve automating information extraction from images and updating geographic information system (GIS) databases. Databases for urban areas are frequently established by analyzing aerial imagery.

Because manual interpretation is very time-consuming, efforts have been made to speed up this process with automatic or semi-automatic procedures. A wide range of techniques and algorithms have been proposed for automatically constructing 2D or 3D building models from satellite and aerial imagery. In this field, Dash, Steinle, Singh, and Bahr (2004) developed a method based on the standard deviation to distinguish between trees and buildings using the height variation at the periphery of the objects present in the data. Based on their studies, Samadzadegan, Azizi, Hahnb, and Lucas (2005) proposed a novel approach for object recognition based on neuro-fuzzy modeling in which structural, textural and spectral information is extracted and integrated into a fuzzy reasoning process. After extraction and connection of edge pixels, Hongjian and Shiqiang (2006) extracted the height of buildings from a sparse laser sample and reconstructed the 3D information for each building. Sohn and Dowman (2007) extracted building tracks automatically from a combination of the IKONOS imagery with pan-sharpened multi-spectral bands and low-sampled airborne laser scanning data. Lafarge, Descombes, Zerubia, and Deseilligny (2008) presented an automatic building extraction method from digital elevation models based on an object approach. In this method, rough approximations of the building footprints are first realized by a method based on marked point processes, and then these rectangular footprints are regularized by improving the connection between the neighboring rectangles and detecting the roof height discontinuities. One of the methods frequently used in building extraction is the snake model. Snakes, or active contour models, were originally introduced by Kass, Witkin, and Terzopoulos (1988). This model, which uses global information about the image contour to obtain a closed or open curve, does not require any prior knowledge about the image. Therefore, it is used widely in many image-processing areas, such as image segmentation, image tracking and 3D reconstruction (Lam & Yan, 1994). The snake model is an energy-minimizing spline guided by external constraint forces and influenced by image forces that pull it toward features such as lines and edges (Shih & Zhang, 2004).

Since the introduction of the original snake model, many studies that extend and improve the snake method have been

* Corresponding author. Tel.: +98 916 3470177; fax: +98 21 88786213.
E-mail address: m_kabolizade@yahoo.com (M. Kabolizade).

0198-9715/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.compenvurbsys.2010.04.006

proposed. Cohen (1991) introduced balloon forces on active contour models. Xu and Prince (1998) introduced gradient vector flow as a new external force on the snake energy. Amini et al. (1990) used dynamic programming to minimize the energy of active contours. Snakes have also been used extensively in various applications. Akgul and Kambhamettu (1999) implemented tracking of the surface of the tongue from ultrasound image sequences under spatial and temporal constraints. Zhang, Stoecker, and Moss (2000) applied active contours to detect the borders of digitized skin tumor images, while Jones and Plassmann (2000) measured the area of leg ulcers using an active contour model with knowledge-based searching. Baumberg and Hogg (1994) tracked an articulated, nonrigid body using active shape models with dynamic filtering. Abe and Matsuzawa (2000) proposed the use of multiple active contours to extract the region of an object. Chalana, Linker, Haynor, and Kim (1996) used multiple active contour models for cardiac boundary detection on echocardiographic sequences. In the field of building extraction using the snake model, Mayungaa, Coleman, and Zhang (2005) developed a semiautomatic method to extract buildings in structured and unstructured urban settlement areas from high-spatial resolution panchromatic imagery. The proposed method used a radial casting algorithm to initialize the snake contours, and the fine measurements of building outlines were carried out using the snake model. Rüther, Martine, and Mtalo (2002) presented a novel approach for semiautomatic building extraction in informal settlement areas from aerial photographs. In the proposed approach, he uses a strategy of delineating buildings by optimizing their approximate building contour position. Approximate building contours are derived automatically by locating elevation blobs in digital surface models. Building extraction is then applied by means of the snake algorithm and the dynamic programming optimization technique. Peng et al. (2005) proposed an improved snake model focusing on building detection from gray-level aerial images with high resolution. His model is implemented with two aspects: the criteria for the selection of the initial seeds and the external energy function. In this paper, a contour detection method based on the snake model is proposed and developed. The proposed method is an automatic method that, in comparison with traditional models, can converge to the true building contours more quickly and more stably even in complex environments.

To define regions of interest using this method, building detection is performed by searching for local maxima in a digital surface model (DSM). This approach is based on the knowledge that buildings are objects of limited size rising from the terrain surface. Within these regions, feature extraction is performed using DSM data in combination with height data. The DSM used to detect buildings can be produced by stereo matching or from airborne laser scanning rangefinder data, although the results of stereo matching are not very good in urban areas. The airborne scanning laser rangefinder can acquire a high density of laser points to generate the DSM of a city (Maas, 1999). Many studies that used aerial and laser scanning data have already been proposed. Rottensteiner, Trinder, Clode, and Kubik (2005) also proposed a new method for the detection of buildings in densely built-up urban areas through the fusion of first- and last-pulse laser scanner data and multispectral images. In this method, land cover is classified into the "building", "tree", "grassland" and "bare soil" classes, with the latter three considered relevant for the subsequent generation of a high-quality digital terrain model (DTM). Building detection is accomplished by first applying a hierarchical rule-based technique for coarse DTM generation based on morphological filtering. Hongjian and Shiqiang (2006) proposed an approach for automatic 3D building reconstruction based on aerial digital images and sparse laser scanning sample points. In their method, a Laplacian sharpening operator and threshold segmentation are used to first extract the edges of the digital image, and then pixel connectivity is used to extract the linear features in the CCD image. A bidirectional projection histogram and line matching are proposed to extract the contours of buildings. The height of the building is determined from sparse laser sample points. Sohn and Dowman (2007) proposed a new approach for automatic extraction of building footprints for a combination of IKONOS imagery and low-sampled airborne laser scanning data. In their method, a laser point cluster in 3D object space is recognized as an isolated building object. As modeling cues, rectilinear lines around building outlines collected in either a data-driven or model-driven manner are integrated to compensate for the weaknesses of both methods. Finally, a full description of building outlines is accomplished by merging convex polygons, which are obtained as a building region that is hierarchically divided by the extracted lines using the Binary Space Partitioning (BSP) tree.

Light detection and ranging (LiDAR) data can be used for two purposes: first, to generate a model depicting the buildings in the area by using height data, and second, to use a snake model for guiding the contours. The model should convey information about the height, shape and other related information about the building. We present an automatic method for extracting buildings based on the snake model with improved external energy using LiDAR data and aerial images in urban areas.

In Section 2, the building extraction method, including the selection of initial points, the traditional snake model and the improved snake model, are described. In Section 3, the experimental results from implementing the new model are presented. In Section 4, the numerical evaluation and the limitations of the method are discussed, and finally, we draw conclusions in Section 5.

2. Building extraction method

The main steps followed in the proposed methodology are given in Fig. 1. The approach consists of three steps to extract the buildings from high-resolution aerial images: (i) generating a digital surface model (DSM) from LiDAR data and a digital elevation model (DEM); (ii) detecting buildings and selecting initial points using normalized digital surface model (NDSM), including generating raised structure hypotheses by global height thresholding of the NDSM; and (iii) improving the snake model, minimizing the energy function and optimizing the approximate building contours.

2.1. Selection of initial points

The overall aim of the detection process was to partition an image into regions where buildings could be constructed. This primary segmentation was carried out to reduce the dimensionality of the search space and consequently to reduce the compu-

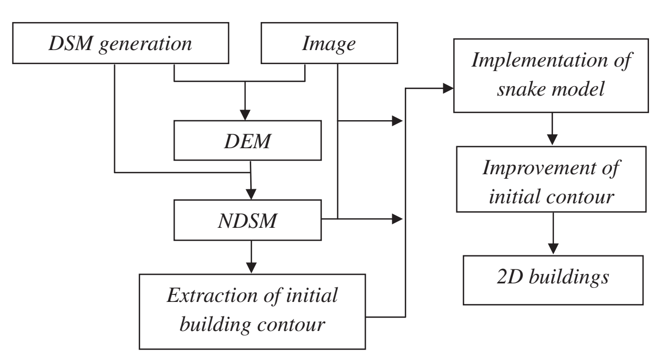

Fig. 1. The main stages of the proposed building extraction method from high-resolution aerial images.