ISPRS Journal of Photogrammetry and Remote Sensing 136 (2018) 26–40

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

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

Recognition of building group patterns in topographic maps based on graph partitioning and random forest

Xianjin He [a], Xinchang Zhang [a,b,c,*], Qinchuan Xin [a,b,*]

[a] School of Geography and Planning, Sun Yat-Sen University, Guangzhou 510275, China
[b] Guangdong Key Laboratory for Urbanization and Geo-simulation, Guangzhou 510275, China
[c] School of Geographical Sciences, Guangzhou University, Guangzhou 510275, China

ARTICLE INFO

Article history:
Received 3 June 2017
Received in revised form 12 September 2017
Accepted 1 December 2017

Keywords:
Building groups
Graph partitioning
Pattern recognition
Machine learning
Cartography

ABSTRACT

Recognition of building group patterns (i.e., the arrangement and form exhibited by a collection of buildings at a given mapping scale) is important to the understanding and modeling of geographic space and is hence essential to a wide range of downstream applications such as map generalization. Most of the existing methods develop rigid rules based on the topographic relationships between building pairs to identify building group patterns and thus their applications are often limited. This study proposes a method to identify a variety of building group patterns that allow for map generalization. The method first identifies building group patterns from potential building clusters based on a machine-learning algorithm and further partitions the building clusters with no recognized patterns based on the graph partitioning method. The proposed method is applied to the datasets of three cities that are representative of the complex urban environment in Southern China. Assessment of the results based on the reference data suggests that the proposed method is able to recognize both regular (e.g., the collinear, curvilinear, and rectangular patterns) and irregular (e.g., the L-shaped, H-shaped, and high-density patterns) building group patterns well, given that the correctness values are consistently nearly 90% and the completeness values are all above 91% for three study areas. The proposed method shows promises in automated recognition of building group patterns that allows for map generalization.

1. Introduction

As common geographical entities in urban areas, buildings are important navigational objects in topographic maps for users. The building group pattern refers to the arrangement and form exhibited by a collection of buildings at a certain scale in the mapping space (Du et al., 2016a, 2016b). Typically, building group patterns can be categorized into regular patterns (e.g., linear, rectangular, and grid-like arrangement) and irregular patterns (e.g., T-shaped, L-shaped, Z-shaped, and H-shaped arrangement) (Anders, 2006; Yang, 2008; Du et al., 2016a, 2016b). Recognition of building group patterns in topographic maps is important to the understanding and modeling of geographic space because the recognized patterns can be simplified while preserving the spatial configuration of the buildings in a scene (Harrie and Weibel, 2007; Haunert, 2012; Mackaness et al., 2014). It hence plays an essential role in a wide variety of downstream applications such as map generalization (i.e., the procedure that utilize transformation operations to solve spatial conflicts of entities and derive smaller-scale maps from large-scale-maps) (McMaster and Shea, 1992; Li et al., 2004; Sandro et al., 2011) and semantic classification of urban structures and functions (Du et al., 2015; Niu et al., 2017).

Detection of building group patterns is challenging because the group patterns are often scale-dependent and vary with building distributions (e.g., the distances, relative orientations, and topological relationships among buildings) (L. Zhang et al., 2013; X. Zhang et al., 2013). A variety of methods have been developed for detecting building group patterns. One method is to first define patterns, such as straight-line, L-shaped pattern, H-shaped pattern and star-shaped pattern, and then match objects with the defined templates (Christophe and Ruas, 2002). Another method considers the proximity of objects and obtains building clusters in each block based on the buffering analysis (Lee, 2004; Basaraner and Selcuk, 2008; Cetinkaya et al., 2015). There are also studies that first identify building pairs with maximum similarities in term of geolocations

* Corresponding authors at: School of Geography and Planning, Sun Yat-Sen University, Guangzhou 510275, China.

E-mail addresses: eeszxc@mail.sysu.edu.cn (X. Zhang), xinqinchuan@gmail.com (Q. Xin).

https://doi.org/10.1016/j.isprsjprs.2017.12.001
0924-2716/

and shapes and then group them based on certain mapping constraints (Li et al., 2004; Yan et al., 2008; L. Zhang et al., 2013; X. Zhang et al., 2013; W. Wang et al., 2015; Y. Wang et al., 2015; Du et al., 2016a, 2016b). Anders (2003) developed a hierarchical clustering method that integrates different neighborhood graphs for automated clustering of natural objects represented by points. Most of the existing studies developed rigid rules for specific patterns based on the topographic relationships among buildings and thus the methods are limited to identifying specific patterns such as linear and grid-like patterns and often involve considerable empirical parameters. Given that most studies have largely paid attention to individual buildings and the relationships (e.g., distance, orientation, shape, size) between each pair of buildings within a group, synthesizing the rigid rules for identifying various building group patterns often lead to spatial conflicts. The combinations of individual buildings that form certain spatial and/or semantic patterns could increase in a factorial manner with the building numbers, making it computationally intensive to derive interpretable clusters in the topographic maps.

Another possible solution to identify building group pattern is to first test building clusters as an integrated object and then further partition building clusters if they do not meet the criteria of building group patterns. Such a process essentially builds a hierarchical structure to organize building objects, and is consistent with our visual system that tends to first recognize potential group patterns from an overall aspect and then determine in details whether the potential patterns meets the mapping criteria such as map generalization rules (Feldman, 2003; Froyen et al., 2015).

The objectives of this study are to (1) develop a framework to identify various building group patterns (e.g., linear, L-shaped, and rectangular patterns) based on graph partitioning, and (2) apply operators to the recognized building group patterns for map generalization.

2. Study materials

Three areas located in the Pearl River Delta in Guangdong Province in Southern China were selected for studying the identification of building group patterns (Fig. 1). The topographic datasets at the scale of 1:2000 were provided by the Guangzhou Urban Planning and Design Survey Research Institute in the Guangdong province in China. Table 1 provides a brief description of the used datasets, including the geolocations of the study areas and the numbers of buildings and blocks. The buildings in ZC are more regular and more evenly distributed than that of HZ and ZS. Building group patterns, such as linear pattern, L-shaped pattern, and rectangular patterns pattern, as exhibited by collections of buildings could be easily identified from the datasets in Fig. 1. These datasets have varied building distributions and building group patterns, providing ideal study cases that are representative of the complex urban environment in southern China.

3. Methodology

According to the spatial distribution of buildings, the building clusters can be classified as regular patterns (e.g., linear and rectangular patterns) and irregular patterns (e.g., T-shaped, L-shaped, Z-shaped, H-shaped, and high-density patterns). This study aims to detect both regular patterns (including collinear, curvilinear, and rectangular patterns) and irregular patterns (including L-shaped, H-shaped, and high-density patterns). The high-density pattern refers to building clusters that include randomly distributed buildings in the urban villages (a unique phenomenon in the Chinese urbanization processes that villages exist in the core areas of well-developed large cities).

Fig. 2 shows the proposed framework for group pattern recognition and map generalization. Our method, named as Graph-based Spatial Clustering Application with Random Forest (GSCARF), involves three major steps with details explained in the following sections.

3.1. Processing of building blocks

The first step is to partition the topographical map of buildings into a series of building blocks and then process each individual building block as an initial disconnected graph for subsequent analysis. A road network is used to separate buildings into different blocks (Regnauld, 2001; Li et al., 2004). A constrained Delaunay

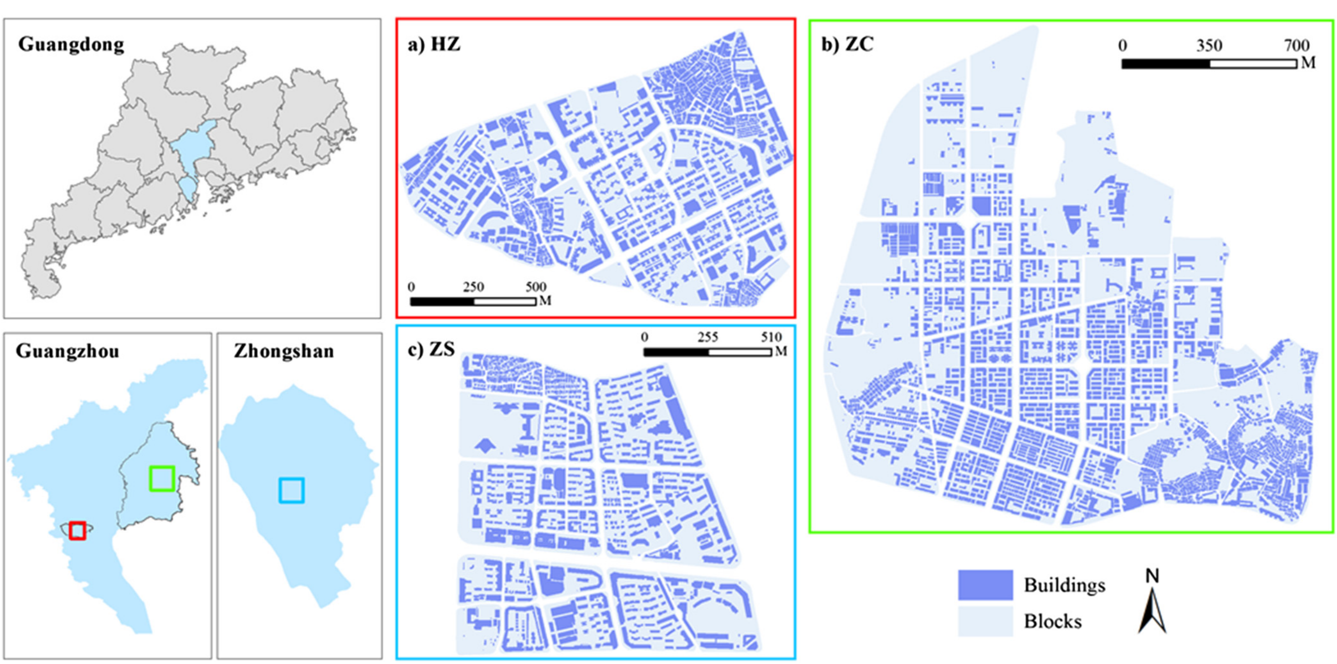

Fig. 1. The experimental datasets are shown for the study area of (a) HZ, (b) ZC, and (c) ZS.