ARTICLE IN PRESS

ISPRS Journal of Photogrammetry and Remote Sensing xxx (2018) xxx–xxx

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

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

Building instance classification using street view images

Jian Kang [a], Marco Körner [b], Yuanyuan Wang [a], Hannes Taubenböck [c], Xiao Xiang Zhu [a,d,*]

[a] Signal Processing in Earth Observation (SiPEO), Technical University of Munich (TUM), 80333 Munich, Germany
[b] Chair of Remote Sensing Technology, Technical University of Munich (TUM), 80333 Munich, Germany
[c] German Remote Sensing Data Center (DFD) (IMF), German Aerospace Center (DLR), 82234 Wessling, Germany
[d] Remote Sensing Technology Institute (IMF), German Aerospace Center (DLR), 82234 Wessling, Germany

ARTICLE INFO

Article history:
Received 9 June 2017
Received in revised form 8 November 2017
Accepted 7 February 2018
Available online xxxx

Keywords:
CNN
Building instance classification
Street view images
OpenStreetMap

ABSTRACT

Land-use classification based on spaceborne or aerial remote sensing images has been extensively studied over the past decades. Such classification is usually a patch-wise or pixel-wise labeling over the whole image. But for many applications, such as urban population density mapping or urban utility planning, a classification map based on individual buildings is much more informative. However, such semantic classification still poses some fundamental challenges, for example, how to retrieve fine boundaries of individual buildings. In this paper, we proposed a general framework for classifying the functionality of individual buildings. The proposed method is based on Convolutional Neural Networks (CNNs) which classify façade structures from street view images, such as Google StreetView, in addition to remote sensing images which usually only show roof structures. Geographic information was utilized to mask out individual buildings, and to associate the corresponding street view images. We created a benchmark dataset which was used for training and evaluating CNNs. In addition, the method was applied to generate building classification maps on both region and city scales of several cities in Canada and the US.

1. Introduction

The classification of land cover from Earth Observation (EO) images in complex urban environments has been a focus in remote sensing over the past decades (Anderson et al., 1976; Pal and Mather, 2003; Yuan et al., 2005; Stefanov et al., 2001; Rodriguez-Galiano et al., 2012; Albert et al., 2017). Beyond, high resolution spaceborne and aerial images are one of the handful information sources for monitoring urban development on large scales.

However, the transfer from land cover to land use in EO-data is complex and relies mostly on the geometry and the appearance of individual buildings and the patterns they group together (Lu and Weng, 2006; Gong et al., 1992; Paola and Schowengerdt, 1995; Pacifici et al., 2009; Khorram et al., 1987; Di et al., 2000; Cheng et al., 2015; Huang et al., 2014, 2015, 2017). The correlation of physical indicators such as building volumes, density or alignment has been used to infer the usage of buildings, e.g. as commercial areas (e.g. Fig. 1(a)), residential areas (e.g. Fig. 1(b)) or industrial areas (e.g. Fig. 1(c)). Nevertheless, such pattern analysis cannot be directly transferable to the classification of individual buildings as we go to a finer level of urban intrinsic scale. For example, Fig. 1(a) shows a commercial area comprised of multiple high-rise buildings. However, the label "commercial area" cannot be assigned to all the building instances within it. As illustrated in Fig. 2, the corresponding street view images show that the commercial area is comprised of a few apartments, office buildings, and one church. This also applies to the example shown in Fig. 1(b) and (c), where both the residential and industrial areas are comprised of buildings with different functionalities. As can be seen, land-use classification at a level of individual buildings is not a trivial task. Usually, such a classification map is only obtainable through city cadastral databases, not accessible or sometimes even not existent. Updating such databases without automatic methods can be very labor intensive. Hence, automatically achieving a building instance-level classification is necessary and can be beneficial for applications related with urban planning. Towards an automatic classification of individual buildings, the challenges are twofold. Firstly, remote sensing images usually only contain roof structures due to their nadir-looking imaging geometry. The visual difference of the roofs between certain building classes, e.g. apartments and office buildings, can be subtle, as an example shown in Fig. 2. Secondly, the extraction of building footprints directly from

* Corresponding author at: Signal Processing in Earth Observation, Technical University of Munich, Arcisstr. 21, 80333 Munich, Germany and Remote Sensing Technology Institute, German Aerospace Center, 82234 Wessling, Germany.

E-mail address: xiao.zhu@dlr.de (X.X. Zhu).

https://doi.org/10.1016/j.isprsjprs.2018.02.006
0924-2716/

Please cite this article in press as: Kang, J., et al. Building instance classification using street view images. ISPRS J. Photogram. Remote Sensing (2018), https://doi.org/10.1016/j.isprsjprs.2018.02.006

Fig. 1. Examples of land-use classification.

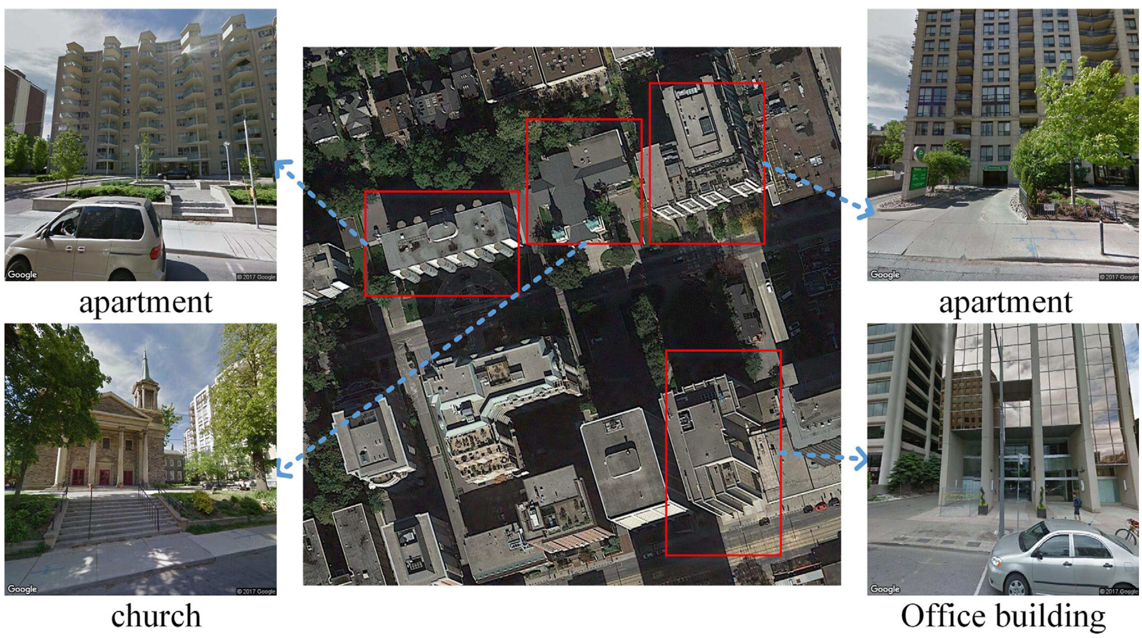

Fig. 2. The commercial land-use area as shown in Fig. 1(a), along with the street view images for some buildings selected by the red rectangles. These buildings do not belong to the same category, even though they are located in the same land-use area. Besides, compared to the roof structures, the information of façade structures displayed in street view images is richer and more sufficient to be used for building instance classification.

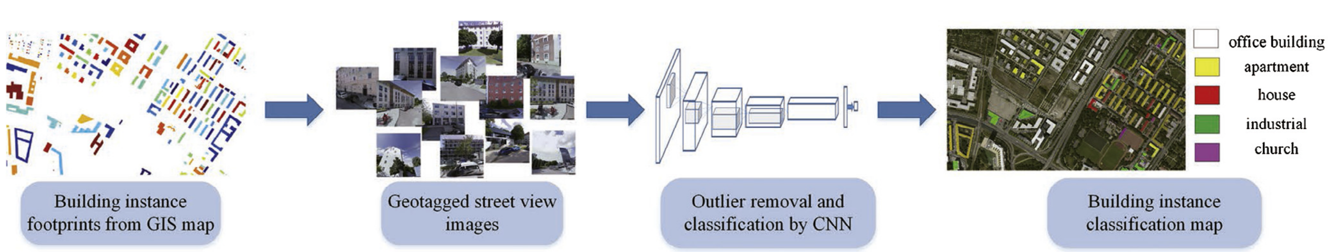

Fig. 3. The proposed workflow for land-use classification at a level of individual buildings.

remote sensing images is still under preliminary research. A clear segmentation of building footprints usually requires height information which comes at an additional cost.

In this paper, we propose a general framework to tackle the abovementioned challenges, which exploits the information extraction from freely available street view images and online geographic maps. Specifically, façade structures shown in online street view images are sufficiently rich for building functionality classification, and the online map services, such as OpenStreetMap (OpenStreetMap, 2017) or Google Maps, can provide the building footprints which can be associated to street view images via their geographic locations. As shown in Fig. 2, the façades displayed in

Please cite this article in press as: Kang, J., et al. Building instance classification using street view images. ISPRS J. Photogram. Remote Sensing (2018), https://doi.org/10.1016/j.isprsjprs.2018.02.006