Building Change Detection Using Semantic Segmentation on Analogue Aerial Photos

Evangelos Maltezos, Charalabos Ioannidis, Anastasios Doulamis, and Nikolaos Doulamis (Greece)

Key words: Informal settlements; Photogrammetry; change detection; buildings; analogue photos; point cloud; convolutional neural networks

SUMMARY

Spatio-temporal urban monitoring in large-scale is critical in various engineering applications. Periodical building change detection is necessary for applications such as urban and rural planning, updating 3D cadastral objects and databases, identification of informal settlements and constructions, 3D city modeling and valuation purposes. Automatic building change detection is a research topic of high interest including several challenges referring to radiometric, geometric and atmospheric corrections, data registration and multi-modal data fusion as well as difficulty in using images of complex building structures derived from different viewpoints. Deep learning techniques have received increased attention for achieving satisfying results in many classification problems. Semantic segmentation is a pixel-wise classification of images by implementing a deep neural network scheme such as Convolutional Neural Networks (CNNs) under a supervised setting. This paper presents an one-shot building change detection procedure using semantic segmentation on scanned analogue aerial photos. An augmented time period feature band vector is firstly created by fusing 3D geospatial information, that is a 3D point cloud extracted from Dense Image Matching (DIM), with the corresponding orthoimage. A small training set for the classes of "new buildings", "unchanged buildings" and "other" is created from the same dataset. The training set and the augmented time period feature band vector are fed as input into a CNN, which is responsible for the semantic segmentation. The "new buildings" and the "unchanged buildings" masks are then processed to eliminate noise taking into account the spatial coherency properties. To verify the applicability, usability and functionality of the employed procedure, two complex and real-life urban study areas in Greece (Keratea and Perissa) with various building structures, pixel resolutions and types of data are selected. The CNN results evaluated via pixel-wise success rates of completeness, correctness and quality, and compared with conventional and other "shallow" learning paradigms, such as Support Vector Machines (SVMs). The derived results show the effectiveness of the proposed deep scheme.