AERIAL PHOTO BUILDING CLASSIFICATION BY STACKING APPEARANCE AND ELEVATION MEASUREMENTS

T. Nguyen[a] S. Kluckner[*a] H. Bischof[a] F. Leberl[a]

[a] Graz University of Technology, Institute for Computer Graphics and Vision, Inffeldgasse 16, 8010, Graz, Austria

Technical Commission VII Symposium 2010

ABSTRACT:

Remote Sensing is trending towards the use of greater detail of its source data, advancing from ever better resolving satellite imagery via decimeter-type aerial photography towards centimeter-level street-side data. It also is taking advantage of an increase in methodological sophistication, greatly supported by rapid progress of the available computing environment. The location awareness of the Internet furthermore demonstrates that large area remote sensing strives for a model of human scale detail. This paper addresses the task of mapping entire urban areas, where objects to be mapped are naturally three dimensional. Specifically we are introducing a novel approach for the segmentation and classification of buildings from aerial images at the level of pixels. Buildings are complex 3D objects which are usually represented by features of different modalities, i.e. visual information and 3D height data. The idea is to treat them in separated processes for learning and then integrate them into a unified model. This aims to exploit the discriminative power of each feature modality and to leverage the performance by fusing the classification potentials at a higher level of the trained model. First, representative features of visual information and height field data are extracted for training discriminative classifiers. We exploit powerful covariance descriptors due to the low-dimensional region representation and the capability to integrate vector-valued cues such as color or texture. Then, a stacked graphical model is constructed for each feature type based on the feature attributes and classifier's outputs. This allows to learn inter-dependencies of modalities and to integrate spatial knowledge efficiently. Finally, the classification confidences from the models are fused together to infer the object class. The proposed system provides a simple, yet efficient way to incorporate visual information and 3D data in a unified model to learn a complex object class. Learning and inference are effective and general that can be applied for many learning tasks and input sources. Experiments have been conducted extensively on real aerial images. Moreover, due to our general formulation the proposed approach also works with satellite images or aligned LIDAR data. An experimental evaluation shows an improvement of our proposed model over several traditional state-of-the-art approaches.

TOPIC: Image processing and pattern recognition
ALTERNATIVE TOPIC: Image processing and pattern recognition