CLASSIFYING SEGMENTED HYPERSPECTRAL DATA FROM URBAN AREAS USING SUPPORT VECTOR MACHINES

Schiefer, S.[1,2], Hostert, P. [1] , Janz A.[3] and Eiden, M. [4]

1. Humboldt-Universität zu Berlin, Geomatics Department, Unter den Linden 6, 10099 Berlin, Germany; Tel. +493020934889, Fax +493020936844; [schiefer, hostert]@hurs.de
2. Center for Remote Sensing of Land Surfaces, University of Bonn, Bonn, Germany
3. Humboldt-Universität zu Berlin, Department of Computer Science, Berlin, Germany; janz@informatik.hu-berlin.de
4. Institute of Chemistry and Dynamics of the Geosphere, ICG-III: Phytosphere, Jülich Research Centre, Jülich, Germany; m.eiden@fz-juelich.de

Urban environments are characterised by great spectral and spatial heterogeneity: objects like houses, courtyards or streets are narrow compared to agricultural or forestal units; the variety of spectrally different, man-made materials is large; surface types change at a high frequency. The classification of remote sensing data from urban environments is thus challenging and requires sophisticated approaches.

One of the main problems in this context is the multi-modal feature space defined by urban surfaces. A class roof, for example, comprises numerous spectrally different materials and illumination conditions.

In addition, different land use classes might show great spectral similarity, e.g. tar roofs and streets. Therefore, data with high spectral resolution, i.e. imaging spectrometry data, is needed for a sufficient distinction of the mentioned materials. However, the high information content leads to large data sets that cannot be handled by some state-of-the art processing techniques.

The concept of image segmentation has proved successful in several contexts. Still, the processing of segmented images is often limited to a small number of algorithms available in proprietary segmentation software. Within this study, segmentation results are reconstructed in a software independent file format. By doing so, a spatially useful data compression is performed that reduces the spectral information of image segments to average values. However, the generalisation reduces the data's spectral variance and decreases the number of training samples; both aspects hinder the use of parametric classifiers like the maximum likelihood classifier.

Support vector machines (SVM) are one example of machine learning algorithms that emerged as promising alternatives to traditional classifiers. They are non-parametric and generalize well even when the number of training samples is small. SVM are capable of modelling complex classes that require non-linear discrimination functions, e.g. multi-modal classes of urban areas.

Data from the Hyperspetral Mapper (HyMap), acquired over the city of Berlin, Germany, in July 2003, are used to map urban surface types in the context of impervious surface mapping. The classes Water, Vegetation, Buildings, Streets and Soils were classified using SVM. All classes comprise a great number of mixed pixels as well as various spectrally different materials.

The present study investigates the influence of the reduced data space of segmented images on classification results. Results show, that good classifications are achieved even at a relatively high degree of generalisation, i.e. high degree of data compression. SVM can be combined with segmented image data due to the software independent file format. They are appropriate for analyses of complex urban environments, in this case the mapping of impervious grounds. The approach is capable of classifying multi-modal classes, even in the reduced data space of segmented images.