USE OF HYPERSPECTRAL AND LASER SCANNING DATA FOR THE CHARACTERIZATION OF SURFACES IN URBAN AREAS

Dirk LEMP, Uwe WEIDNER

Institute of Photogrammetry and Remote Sensing, University Karlsruhe, Englerstraße 7, 76128 Karlsruhe, Germany
{lemp,weidner}@ipf.uni-karlsruhe.de

KEY WORDS: Hyper spectral, LIDAR, Reconstruction, Classification, Urban

ABSTRACT

A recent project of the Engler-Bunte-Institute (EBI), chair of water chemistry, and the Institute of Photogrammetry and Remote Sensing (IPF) aims at the quantitative assessment of pollutants on urban surfaces by chemical analysis and image processing methods. The motivation of this project is the fact that nowadays a better part of the rain water from sealed urban surfaces is treated in sewage plants, although this might not be necessary, because the load of pollutants of the first flush is much higher than in the following run-off. Therefore, the dimensioning of sewage systems may be adopted to this observation and costs may be reduced. In the project, the research focus of EBI is the chemical analysis of washed off pollutants and modelling of the resulting pollution (run-off), whereas the research at IPF deals with the characterization of urban surfaces, namely their geometry (slope, exposition, size) and their surface material. For this purpose two different types of data are used: hyperspectral and laser scanning data with 4 and $1\ m$ planimetric resolution respectively. We combine these data sets of high geometric and spectral resolution to create a detailed map of sealed urban surfaces. The laser scanning data will not only be used to derive geometric properties of the surfaces, but also to improve the classification of materials as it helps for the discrimination of roof and ground surface materials with similar spectra. The paper will present first results of data analysis, which will be focussed on roof surfaces in a first step.

1 INTRODUCTION

In the year 2000, the European Union implemented the water framework directive. This regulation oblige every member state to review the impact of human activity on the status of surface waters and on groundwater. In a recent research project we focus on a small, but nevertheless important topic in this context: the assessment of pollutants on urban surfaces and their impact on the pollution load. Thus, one aim of the project is not only to derive information on the amount of sealed surfaces in an urban area (cf. (Butz and Fuchs, 2003)), but also to derive a detailed surface material map. Therefore, the work package consists of five subtopics – chemical measurements for the characterization of the chemical processes on reference roof surfaces, determination of surface geometry, classification of surface materials, modelling of the resultant pollution, and model verification. In this paper, we describe our work on two of these subtopics, namely the information derivation of the surface characteristcs, i.e. geometric and material properties.

Urban areas are characterized by their complex geometric structure and their heterogenity concerning the occuring surface materials. The appearance of surface patches' materials in the data is influenced by the acquisition and object geometry. Furthermore, the age of the material and environmental conditions, e.g. by weathering and humidity, also have impact on their appearance. All these facts lead to the necessity of high resolution input data to solve the tasks – high resolution with respect to the geometric resolution, but also to the spectral resolution in order to discriminate the various surface materials. Therefore, we combine data derived from laser scanning, which provides the necessary geometric information, and hyperspectral data for the classification of surface materials.

In the following, we give a short overview on related work dealing with the combination of laser scanning and hyperspectral data. Section 3 introduces the input data. Our approach for the characterization of surfaces in urban areas is presented in Section 4 focussing on roof surfaces in a first step, followed by a summary of recent results in Section 5 and the conclusions.

2 RELATED WORK

Up to now, the two data types were often used exclusively, either to derive the geometry based on laser scanning data (cf. (Vögtle and Steinle, 2003)) or to derive material maps based on hyperspectral data (cf. (Heiden et al., 2001)). The improvement of reconstruction from laser data by additional image information is discussed, but mainly to reject vegetation areas. (Gamba and Houshmand, 2000) use hyperspectral data (AVIRIS) in order to improve reconstruction results based on IFSAR, namely to mask vegetation areas, but the used data has only limited resolution. (Madhok and Landgrebe, 1999) integrate DSM information in order to improve the results of hyperspectral classification based on HYDICE data. In their research the DSM, derived from aerial imagery, is applied for the discrimination of roofs and ground surfaces. The materials may have a similar spectrum, but they can be discriminated based on the height information. (Simental et al., 2003) combine hyperspectral (HyMap) and laser data to derive a mobility and trafficability map in an open area, thus the requirements seem to be less strong than in our application.

The approach of (Bochow et al., 2003) is the closest related work to our approach. They use a normalized Digital Surface Model (nDSM) – the difference between a DSM and a DTM approximating the ground surface – with a planimetric resolution of $0.5\ m$ and hyperspectral data taken by the airborne DAIS 7915 sensor and interpolated to $0.5\ m$. The surface model was derived from HRSC-A data and the non-building areas were masked by building outlines from digital cadastral data. They investigated two approaches for the fusion of the data - first, a fusion on signal level and applying Spectral Angle Mapper (SAM) for classification based on 16 channels of a minimum-noise-transformed data set, and second, on a decision level using a binary decision tree.

Our approach differs from the above with respect to the input data, in particular the laser scanning data. We use eCognition, which allows a hierchical classification and introduction of knowledge by using the different information sources for different decisions within a fuzzy classification scheme. Details are given in Section 4.

3 DATA

For the characterization of urban surfaces with respect to their geometry and their materials, two different data sets are combined: a DSM and hyperspectral data.

The DSM was acquired in March, 2002, with the TopoSys system using the first (cf. Fig. 1 and 3) and the last pulse modes. For ease of use within different software packages, $1\ m \times 1\ m$ raster data sets were generated. These data sets differ not only concerning the objects included, but also showing systematic effects: surface patches appear smoother and building footprints are systematically smaller in the last pulse data. The impact of these differences on the analysis will be discussed in Section 5.

The hyperspectral data was acquired in July, 2003, with the HyMap sensor during the HyEurope campaign organized by the DLR (German Aerospace Center). Figure 2 displays a band combination ranging from the visible to near infrared spectrum (cf. Fig. 3). The white line indicates the central campus area. The data was preprocessed (atmospheric corrections, geocoding) by the DLR, Oberpfaffenhofen, using the DSM. The original data has a ground resolution of $4\ m \times 4\ m$. In order to use the data in combination with the DSM, the data was resampled to a resolution of $1\ m \times 1\ m$. We applied different standard techniques like (Dell'Aqua and Gamba, 2003) and their impact on the results of our approach will also be discussed in Section 5.

Dimensionality of hyperspectral data is always of interest. In order to get a first insight, we tried different techniques for band reduction. We applied standard principle component analysis (PCA), minimum-noise-fraction transformation (MNF), and manual selection of bands based on the spectra of surface materials (Fig. 4). The same training sites were used to analyse the class separability using the Batthacharyya distance. For the PCA and MNF data one band after the other were included. Already 12 MNF-bands and 15 PCA-bands lead to a high separabilty based on this distance measure.

Figure 1: nDSM from laser data (first pulse mode)

Figure 2: Hyperspectral data (RGB=25/15/10)

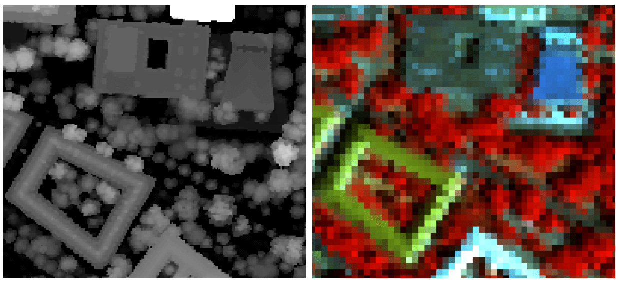

Figure 3: nEnlargement of subset: nDSM from first pulse laser scanning data (left), HyMap data RGB=25/15/5 (right)

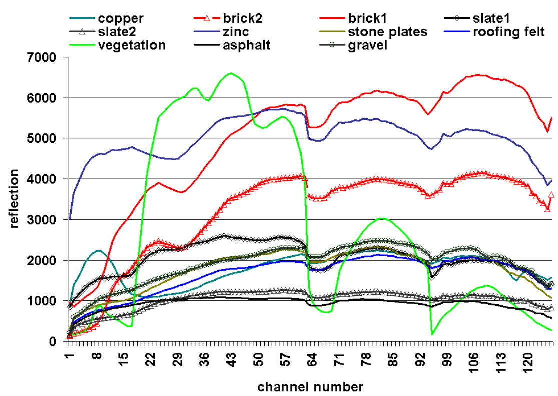

Figure 4: Spectra of selected surface materials

4 APPROACH FOR DATA ANALYSIS

Our approach for the characterization of urban surfaces is based on the analysis of laser scanning and hyperspectral data as depicted in Fig. 5. The geometry of surface patches is derived using a DSM from laser scanning, whereas the surface material information is obtained from both, laser scanning and hyperspectral data. Of course, the hyperspectral data is the main source for the surface material classification, but the used surface material also restricts the geometry or vice versa, the geometry restricts the use of materials. Table 1 shows some examples of roof surface characteristics, grouped with respect to similar spectra, and also indicating qualitatively the surface geometry. Therefore, this information can be used as additional clue within the classification in case the spectral characteristics of different surface materials are almost similar (see Fig. 4).

The main part of our analysis is performed using the software package eCognition. In this software the first step of data analysis is a segmentation, followed by classification of the segments. Therefore, the quality of segmentation is crucial for the quality of classification. In the following, we will describe both steps in detail using a subset of the data as example (white dashed line in Fig. 3).

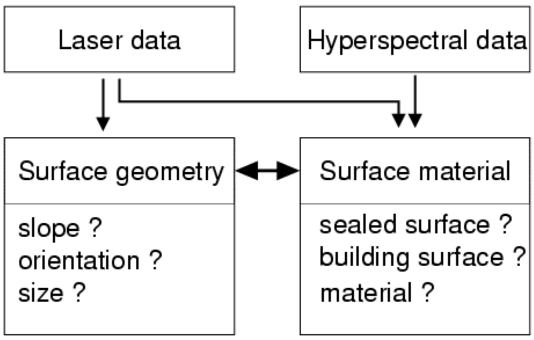

Figure 5: Flow chart of approach

Material	Geometry flat	Geometry sloped	Remarks
Brick	-	+	
Slate	-	+	spectrum similar to stone plates, gravel, roofing felt
Stone plates	+	-	spectrum similar to slate, gravel, roofing felt
Gravel	+	-	spectrum similar to stone plates, slate, roofing felt
Roofing felt	+	+	spectrum similar to stone plates, slate, gravel
Copper	+	+	both possible; sometimes just facing at roofs' outlines with other material like gravel for the main part;
Zinc	+	+	see remarks for copper
Gras	+	+	limited slope

Table 1: Examples of roof surface characteristics

4.1 Segmentation

The segmentation procedure within the eCognition software is based on a region growing algorithm. The criterion for the growing combines three different quantities: the homogeneity of the segment, the shape of the segment measured by its compactness, and the smoothness of its boundary. The homogeneity of the segment takes the deviations from the mean of each channel used for segmentation into account. Thus, the underlying model assumes constant values for each segment's channel, which is only adequate when dealing with flat roofs, but not when dealing with roofs consisting of planar faces, which is our assumed model, and using the laser scanning data as main information for the segmentation. Aware of this problem, we nevertheless tried the segmentation procedure of eCognition. Examples of these segmentations are given in Fig. 6 and 7. For these segmentations first and last pulse data and a NDVI (channels 25 and 15 of the HyMap data) are used. Emphasis was on the geometry data (each channel with weight 4), and less on the NDVI data (weight 1). The segmentations are based on two different scale parameters. A visual inspection of the results indicates what was already expected: The gable roof of a building in the lower left corner (cf. Fig. 8) was segmented into several slight elongated segments in the main roof directions, just approximating the sloped surface by segments with constant heights - independent from the choice of scale parameter. In case of flat roofs, e.g. building the upper middle, the segmentation resulted in reasonable segments, when considering, that there are smaller extensions on this roof (cf. Fig. 8).

Instead of the segmentation by eCognition, our segmentation procedure for laser scanning data searches for planar faces. It follows the region growing principle taking the deviation from a plane in 3D into account. Details of the algorithm are given in (Vögtle and Steinle, 2000). Fig. 9 shows the result of the algorithm for the subset based on the last pulse laser scanning data, thus only the geometry is taken into account during segmentation. Parameters were set to include smaller roof extension in the surrounding larger surface patch. The use of geometric data only may lead to problems, when one planar roof surface patch consists of areas with different surface materials. In order to overcome this drawback, the segmentation may be introduced into eCognition and a second step of segmentation using the spectral data to split up the initial segments may be performed if needed. In this case, segmentation and classification are closely related, because those channels carrying the information for classification should also be used for the segmentation. For the classification described in the next section, we used the results of the eCognition segmentation with scale parameter 50 shown in Fig. 6 and the initial segments without refinement of our segmentation (Fig. 9).

4.2 Classification

Fig. 4 displays example spectra of materials to be classified. A closer look reveals the following:

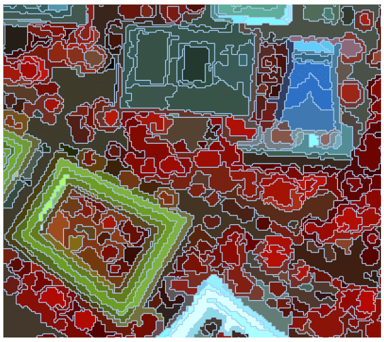

Figure 6: Segmentation (eCognition, scale parameter 50)

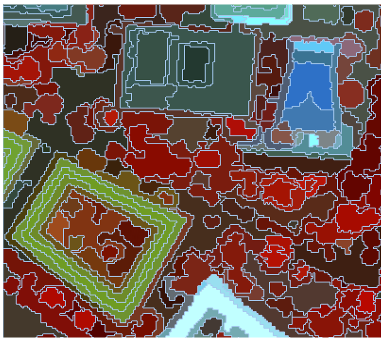

Figure 7: Segmentation (eCognition, scale parameter 75)

Figure 8: Aerial image of buildings

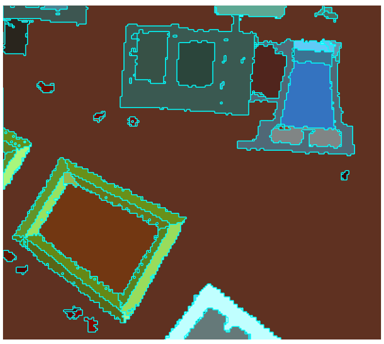

Figure 9: Segmentation (roof planes)

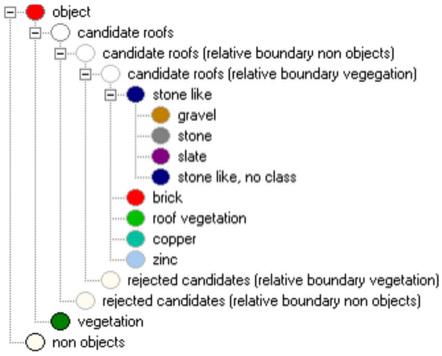

Figure 10: Class hierarchy

- Some materials show a significant different spectrum than the others, e.g. zinc and copper.
- Some spectra of different materials are quite similar, e.g. stone plates and gravel.
- Spectra of same material differ significantly due to the surface orientation in relation to the sun angle/ illumination, e.g. brick or slate.

Therefore, the main tasks are (1) to find specific characteristics of the spectra and select channels from the hyperspectral data for the classification, and (2) to find quantities derived from the available channels, which reduce the influence of illumination. Furthermore, those materials showing a significant spectrum should be classified first, thus leading to a hierarchy in classification. The hierarchy we used is depicted in Fig. 10. First, we classify *objects* and *non objects* using the height information from laser scanning (first and last pulse). In a second step we derive a set of candidate roofs to be classified, by removing vegetation areas from the *objects* applying an NDVI (channel 25 and 15 of the HyMap-data) and smaller segments based on their size and their neighbourhood relations to segments of the classes *non object* and *vegetation*. Thus, this classification procedure may in principal also be applied, if only a nDSM from first pulse data or derived from other sensor data is available. The roof segments are now classified according to their material. For this purpose, we first have to define membership functions for each class and feature to be used, starting with those material classes with the most significant spectral differences to other materials. Zinc has high reflection values in the first channels and show some characteristic slopes, but these features seem to be different for new and and older zinc roofs. Therefore, the fuzzy *or(max)* is used to compute the membership function value from the values of each feature. The spectrum of copper has a significant decrease from channel 8 to 20. Brick shows an increase in the spectrum from the first channels to the last, which seems in our case to be independent from the age of the material. Slate, stone plates and gravel are quite similar with respect to their spectra, but show differences in channel combinations, although not as significant as decreases or increases of the spectra of the other materials above. Therefore, we tried different approaches for the computation of the class membership values based on *and(min)*, *or(max)*, and *mean(arith.)* and introduced also a class *stone like*, if no class of *gravel*, *stone*, or *slate* is assigned.

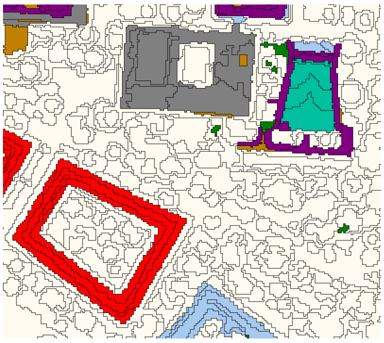

Figure 11: Classification (eCognition, scale parameter 50)

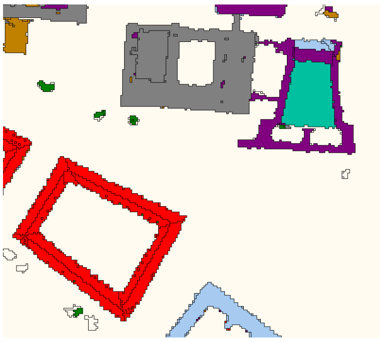

Figure 12: Classification (roof planes)

Fig. 11 and 12 show the results for the classification based on the eCognition segmentation and our segmentation respectively. The subset is also shown in Fig. 8 for comparision. The roof of the surrounding hallway of building 30.21 (upper right corner) is made of zinc, but classified as slate. This seems to be due to the resolution of the hyperspectral data, because the width of the hallway is approximately 2 m, thus only half the original pixel size. For the examples above, the hyperspectral data was resampled using nearest-neighbour interpolation. We will address this point also in the next section.

5 RESULTS

In this section we will present and discuss results of our approach. For this purpose we will focus on the central campus area (white line in Fig. 2), because for this area some reference data already exists, namely a database of buildings with information about their roof materials.

Fig. 13 displays the result of surface material classification based on the segmentation by eCognition. For the classification we used hyperspectral data resampled to 1 m using nearest-neighbour interpolation. We furthermore used first and last pulse laser scanning data. First pulse data includes more details, last pulse data already generalizes the result, because smaller details are not included. The shown roof segments represent those, which are also included in the last pulse data. The membership values of the classes *gravel*, *stone*, and *slate* are computed using the fuzzy *or (max)*. A visual check of the results indicates that the classification delivered reasonable results. Problems arise at

Figure 13: Classification (OR, eCognition)

Figure 14: Stability (OR, eCognition)

Figure 15: Classification (AND, eCognition)

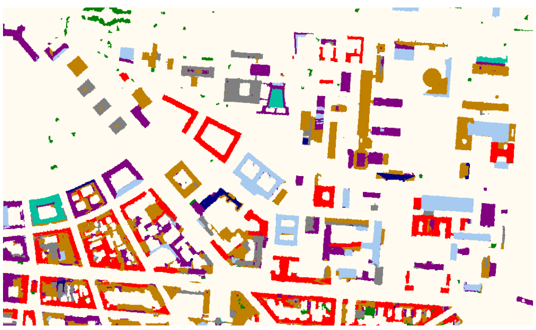

Figure 16: Classification (OR, IPF)

borders of buildings and for smaller segments. Tests using bilinear or cubic interpolation were performed, but show only minor changes of the results. The main problem is the separability of the classes *gravel*, *stone,* and *slate,* which also becomes obvious checking the stability of the classification results (cf. Fig. 14). Most of the roof segments with unstable result – i.e. second best classification result has only small difference in its membership value compared to the best – belong to the above mentioned classes. These segments are shown in red. In case the fuzzy *or(max)* is used, already one feature with high membership value is sufficient for classification. If we use the fuzzy *and(min)*, all feature membership values have to be high for a class to be selected. Fig. 15 shows the results for fuzzy *and(min)*. A number of segments are only classified as *stone like* with higher classification stability.

Fig. 16 shows the result of classification based on our segmentation using last pulse laser scanning data as input. A visual comparision with the result in Fig. 13 – both based on fuzzy *or(max)* – does not show large differences in classification. Differences occur in case the material in a planar patch changes or two roof surfaces are segmented as one segment, because the change in geometry is only small (only small height differences, smooth transition from one roof plane to another), thus indicating that a refinement by using the spectral information as described in Section 4.1 is mandatory.

Up to now, no geometric information has been used for the classification of the roof surface materials. We expect that introduction of gradient information as additional clue may help to discern at least *slate* from *gravel* and *stone*. First tests based on gradients directly derived from the laser scanning data indicate that gradient information should not be extracted directly from the laser scanning data, but from roof planes or segments to give reasonable results.

6 CONCLUSIONS

In this contribution we presented our approach for the characterization of urban surfaces, focussing in a first step on roof surfaces. Input data are laser scanning and hyperspectral data, which are analysed using the software package eCognition and our own software for the segmentation of laser scanning data. First results are presented, which show in principle the feasibility of our approach. The main problems with respect to classification of surface materials are the variability of the materials on one hand and the similarity of some materials' spectra on the other hand. A classification based only on the hyperspectral data is difficult, although the data provides high spectral resolution. We therefore intend to include geometric properties, namely the slope of roofs, into our approach. Furthermore, a quantitative evaluation of the results is necessary. Up to now our reference data is only coarsely related to the buildings and has to be improved to serve as reference data for single roof segments. The ongoing research by the Engler-Bunte-Institute, Chair of water chemistry, on the chemical processes on roof surfaces, will also influence our work, because this research will indicate, which surface materials have to be discerned and which may be grouped with respect to the resultant pollution, thus the requirements on the classification may still change.

ACKNOWLEDGEMENTS

The project is funded by Ministerium für Wissenschaft, Forschung und Kunst Baden-Württemberg – Forschungsschwerpunktprogramm Kapitel 1423 Titelgruppe 74, *Quantitative Assessment of Pollutants on Urban Surfaces by Chemical Analysis and Image Processing Methods*. The authors also would like to thank FGAN-FOM, Ettlingen, for their co-financing of the hyperspectral data aquisition and Eberhard Steinle, who provided the software for the segmentation of roof segments.

REFERENCES

Bochow, M., Greiwe, A. and Ehlers, M., 2003. Ein Prozessmodell zur Analyse von Hyperspektraldaten in urbanen Gebieten. In: Vorträge 23. Wissenschaftlich-Technische Jahrestagung der DGPF, Bochum, pp. 255 – 264.

Butz, J. and Fuchs, S., 2003. Estimation of sealed surfaces in urban areas and the impact on calculated annual pollution loads due to combined sewer overflows. In: C. Jürgens (ed.), Remote Sensing of Urban Areas, IAPRS, XXXiV-7/W9, pp. 35 – 40.

Dell'Aqua, F. and Gamba, P., 2003. Using image magnification techniques to improve classification of hyperspectral data. In: IGARSS 2003, Toulouse. CD-ROM.

Gamba, P. and Houshmand, B., 2000. Hyperspectral and IFSAR data for 3d urban characterization. In: IGARSS 2000, Honolulu. CD-ROM.

Heiden, U., Roessner, S. and Segl, K., 2001. Potential of hyperspectral hymap data for material oriented identification of urban surfaces. In: Remote Sensing of Urban Areas, Regensburger Geographische Schriften, Vol. 35, pp. 69 – 77.

Madhok, V. and Landgrebe, D., 1999. Supplementing hyperspectral data with digital elevation. In: IGARSS 1999, Hamburg. CD-ROM.

Simental, E., Ragsdale, D., Bosch, E., Dodge, R. and Pazak, R., 2003. Hyperspectral dimension reduction and elevation data for supervised image classification. In: 14th ASPRS Conference, Anchorage, AK.

Vögtle, T. and Steinle, E., 2000. 3d modelling of buildings using laser scanning and spectral information. In: IAPRS, Vol. 33, Part B3, pp. 927 – 934.

Vögtle, T. and Steinle, E., 2003. On the quality of object classification and automated building modelling based on laserscanning data. In: ISPRS WG III/3 Workshop - 3D Reconstruction from Airborne Laserscanner and InSAR Data. CD-ROM.