'Photogrammetric Week '99' D. Fritsch & R. Spiller, Eds., Wichmann Verlag, Heidelberg, 1999.

Combining Multiple Data Sources for Urban Data Acquisition

NORBERT HAALA, Stuttgart

ABSTRACT

One of the most promising approaches to solve the challenging task of thematic mapping in urban environments is the integration of multiple data sources to benefit from the complementary types of available information. Within the paper the combination of Digital Surface Models (DSM) resulting from airborne laser scanning and colour aerial imagery for landuse classification in urban environments is described. The laser DSM is used to obtain information on the local height above the terrain surface for each pixel. This information can be applied in order to separate objects higher than the terrain like buildings and trees from objects at the terrain level like streets and grass-covered areas. The basic idea of the proposed algorithm is to combine the geometric information with additional multispectral information provided from colour aerial imagery. Both types of information are integrated in a pixel-based classification, where the information on the local height above the terrain is applied as an additional channel together with the spectral channels. It is demonstrated that by this approach the classification of urban scenes can be improved considerably compared to a scene labelling, where only one type of data is applied.

1. INTRODUCTION

Spectral information has been widely used as a data source for thematic mapping applications. Surface material information can be derived by traditional classification techniques from multispectral imagery and can for example be used for mapping of man-made structures and natural features in complex urban scenes, which is the main purpose of the work described in this paper. During the past years numerous classification algorithms have been developed for thematic mapping applications. (Lillesand and Kiefer (1994)). Still, the main problem during thematic mapping of man-made structures and natural features is the limited accuracy and reliability of the results as well as the frequently arising difficulties to discriminate a sufficient number of object categories. A common goal during data acquisition in built-up areas is the detection of objects like streets and buildings. However, this can be difficult if only spectral information is used, since for some areas roofs and streets are build of very similar material. This complicates or even prevents the discrimination of these objects due to their similar reflectance. The same problem can arise, if trees and grass-covered areas have to be differentiated.

Traditionally, multispectral data has been the dominating source of information for landuse classification. Usually, the input data for the classification are multispectral images. Optionally, textural patterns, which are derived from the original spectral data can be included. Up to now auxiliary information on the surface topography has been mainly used for the geometric correction of the spectral data during the generation of ortho images. Additionally, surface topography information has been used for the radiometric correction of the multispectral imagery. This can be necessary especially in hilly or mountainous terrain, since spectral reflectance values are influenced by the inclination and orientation of the sensed terrain surface. In contrast to this, in our approach height data is not only applied for correction of the multispectral data, but is integrated as a complementary data source of equally important information.

During data acquisition in urban environments, trees and buildings are object classes of major interest since these objects are very relevant for many applications like visualisations or simulations on the propagation of electro-magnetic radiation. In order to separate these types of objects from their surrounding, information on the local height above terrain can be very valuable since trees and

'Photogrammetric Week '99' D. Fritsch & R. Spiller, Eds., Wichmann Verlag, Heidelberg, 1999.

buildings are higher than their surrounding, whereas other objects of interest like streets and grass-covered areas are at the terrain level. The required information can be derived at least approximately from a Digital Surface Model (DSM), which for our application is acquired by airborne laser scanning. Based on this DSM the local height above terrain of the visible surface is derived. The result consists of all objects rising from the terrain approximately put on a plane. This surface is combined with the multispectral bands during a standard classification in order to improve the extraction of the required object classes.
Since the use of geometric information during multispectral classification is the main purpose of the algorithm, the acquisition, processing and analysis of the height data is of major interest. For this reason these steps will be described in detail in the following section.

2. GEOMETRIC INFORMATION

The geometric information, which is used during the classification step has to be extracted from a so-called Digital Surface Model (DSM). This DSM can be obtained by automatic image matching algorithms applying stereo aerial imagery or can be directly captured by airborne scanning laser systems. These airborne laser scanning systems are capable to measure three-dimensional points on the terrain surface dense and well-distributed. For that purpose, similar to a tachymetric data acquisition the coordinates of terrain points are determined by polar measurement. For point determination the run-time of a laser pulse reflected at the ground is used to calculate the distance between the sensor and the laser footprint, whereas position and orientation of the sensor is determined by an integrated GPS/INS system. By current systems terrain points can be measured at an accuracy in the order of 0.3 m (Lohr and Schaller (1999)). Depending on the sensor system and the platform, which can be either helicopter or fixed wing, measurement densities up to several points per square meter can be obtained. An example of a DSM acquired by laser scanning at a point density of one point each 1 x 1 m^2 is shown in Figure 1. In order to improve the 3D visualisation of the data, the ortho image has been wrapped over the DSM surface.

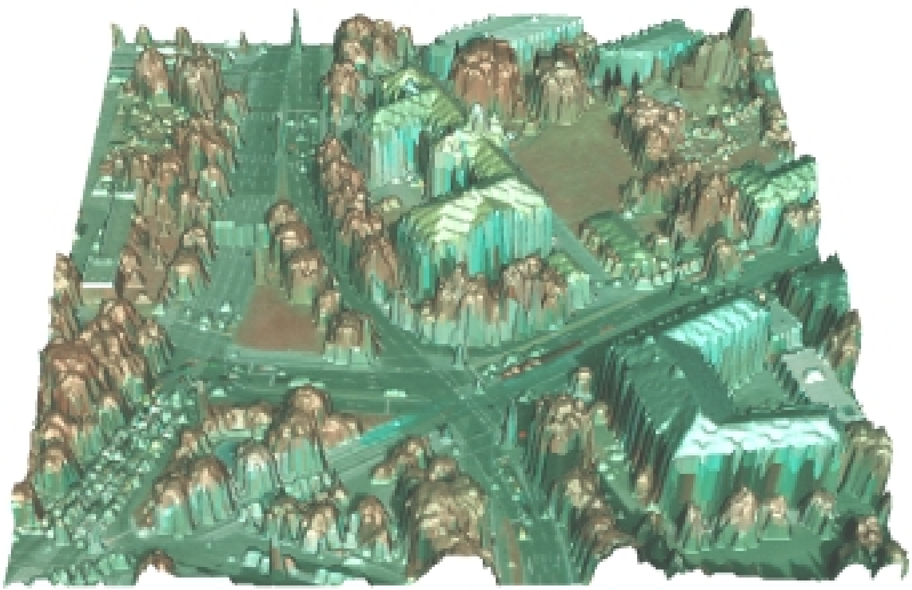


Figure 1: 3D view of test area, height data obtained from laser scanning.

'Photogrammetric Week '99' D. Fritsch & R. Spiller, Eds., Wichmann Verlag, Heidelberg, 1999.

In principle image matching can alternatively be applied for height data acquisition. The main advantage of this approach is that the same sensor, namely an airborne camera, can be used to provide the geometric and radiometric information, which is required in our approach. Image matching techniques have become standard tools for three-dimensional surface acquisition in open terrain. Still they suffer from problems in built-up areas due to occlusions and height discontinuities. In these areas the DSM quality mainly depends on the presence of texture at roof regions and on the amount of contrast between roof and terrain surface. This results in considerable differences of DSM quality especially at roof regions, even in the same image pair.

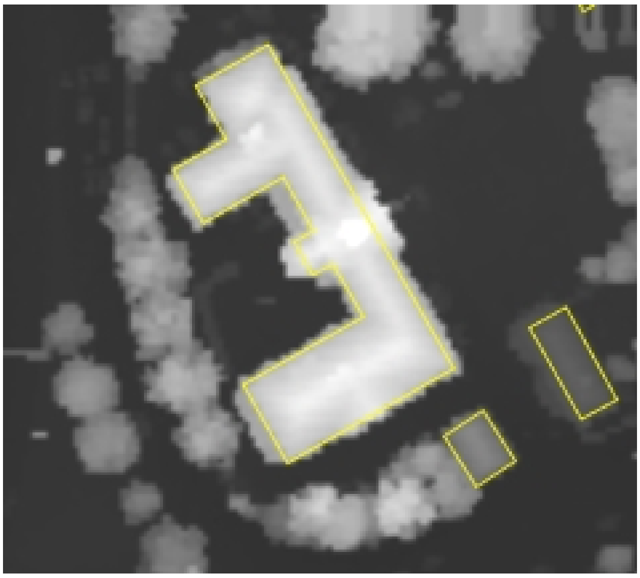


Figure 2: Height data provided from laser scanning.

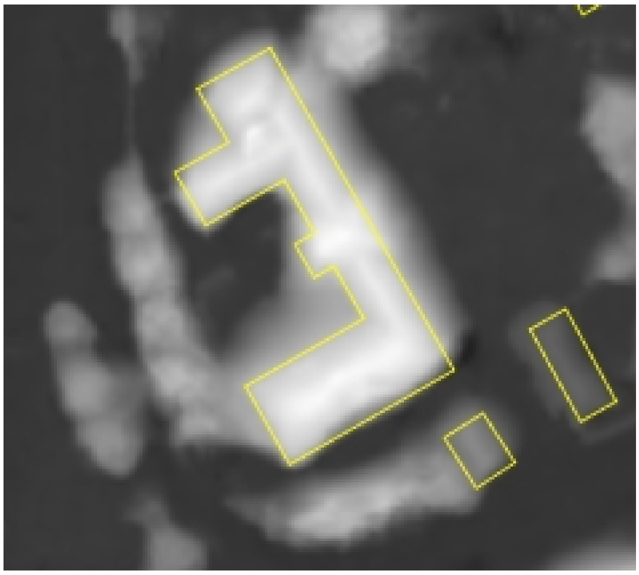


Figure 3: Height data provided from stereo image matching.

An example for the different quality of height data in urban areas is given in Figure 2 and Figure 3. Figure 2 shows a grayvalue representation of height data acquired by laser scanning, Figure 3 gives the corresponding result obtained from stereo image matching. For better visualization the ground plan polygons of the buildings, which were digitized from a 1 : 5000 scale map are overlaid. Whereas for the DSM from image matching very smooth transitions between roof and terrain surface are visible for some areas, these breaklines are defined more sharply in the DSM from laser scanning.
A lot of effort has been spent on software development to overcome the deficiencies of stereo image matching in build-up areas for example by integrating knowledge on predefined breaklines or by using multiple images (Schlüter (1998)). An additional improvement can be expected due to the increasing of radiometric image quality if digital aerial cameras are applied. Still the direct height measurement by airborne laser scanners usually provides DSM data of higher and more homogeneous quality especially in urban areas.

2.1. Preparation of Height Data

In the example data sets depicted above it is clearly visible that DSM not only represent the terrain surface like Digital Terrain Models (DTM), but also contain buildings and other objects like trees, which are higher than their surroundings. To make the information on these objects accessible the so-called normalised DSM, i.e. the difference between DSM and DTM has to be calculated as the

'Photogrammetric Week '99' D. Fritsch & R. Spiller, Eds., Wichmann Verlag, Heidelberg, 1999.

first step. This results in a representation of all objects rising from the terrain approximately put on a plane. The required DTM can be derived from the measured DSM by mathematical grey scale morphology like it is suggested by Weidner and Förstner (1995). In their approach the DSM is processed by a morphological erosion, which is followed by a morphological dilation. The combination of erosion and dilation results in an opening of the DSM surface, eliminating all local maxima in height of a predefined size. This predefined size is for example based on the expected maximum extend of a building and is used a priori in order to define the size of the morphological operator to be applied. After morphological processing the resulting surface roughly represents the terrain surface, i.e. an approximate DTM has been created.
In principle also available 2D GIS information can be used to extract terrain points from the DSM by applying the position of objects known to be at terrain level. An example is the use of the German topographic cartographic database ATKIS, where streets are represented by linear objects. Height values at these regions can be extracted from the measured DSM and utilized to generate the required DTM by interpolating between these areas.

2.2. Object Detection

DSM have already been used for the automatic detection of buildings to trigger their subsequent geometric reconstruction from stereo image data (Baltsavias, Mason and Stallmann (1995)). Even though objects rising from the terrain can be detected quite well from the height data, the discrimination between buildings and trees can be difficult, if only simple criteria like region size or shape are considered. A possible approach is to use the roughness of the DSM surface measured by differential geometric quantities as an additional criterion for the discrimination of buildings and vegetation (Brunn and Weidner (1997)), (Maas (1999)). However, due to the restriction to surface geometry, the number of object types, which can be discriminated within a DSM is limited. Additionally, the use of surface roughness or height texture requires the availability of relatively dense DSM data.
Another possibility to improve object detection based on DSM resulting from airborne laser scanning is the further analysis of the reflected laser beam. Laser scanning systems can be separated into continuous wave or pulsed laser systems. If a pulsed laser system is applied, multiple reflections will occur during the acquisition of trees. As depicted in Figure 4, during measurement of trees a certain percentage of the laser footprint will be reflected by the branches and leaves of the tree. Other parts will penetrate the foliage and will be finally reflected by the terrain surface. For this reason the top of the tree refers to the first response of the laser pulse, which is received by the laser sensor, while the last pulse usually refers to the terrain surface.

'Photogrammetric Week '99' D. Fritsch & R. Spiller, Eds., Wichmann Verlag, Heidelberg, 1999.

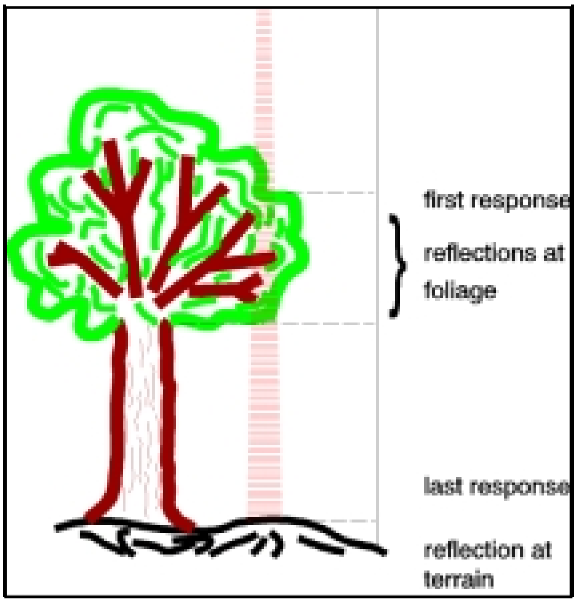


Figure 4: Reflection of laser pulses at trees.

If the laser system is capable to discriminate multiple responses, they can be utilized in order to detect trees and buildings. Figure 5 gives a gray value representation of a normalized laser DSM. The original DSM, which was already depicted as 3D visualization in Figure 1 is based on the first pulse measurement. For this reason both trees and buildings are visible. Figure 6 gives the corresponding result if a DSM based the last pulse measurement is applied. In this example only the buildings are visible. Hence the difference between both normalized DSM can be used for the detection of tree regions.

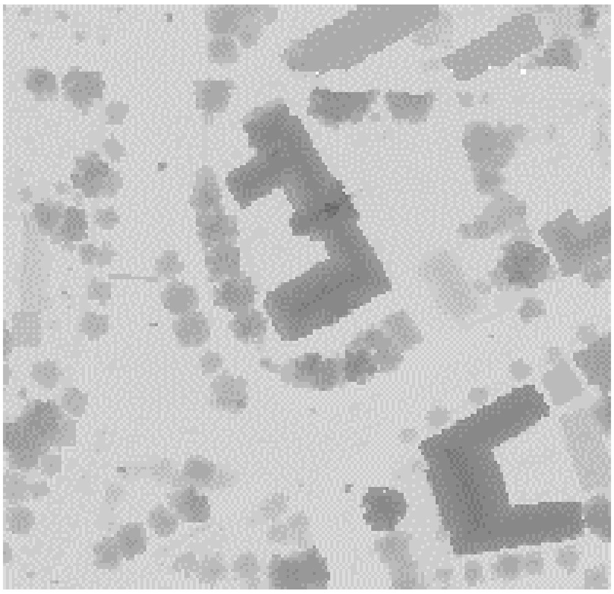


Figure 5: Grey value representation of DSM derived from first pulse measurement.

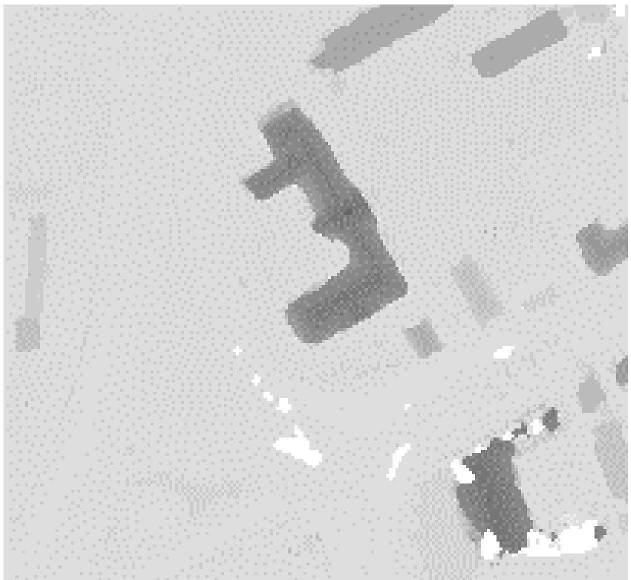


Figure 6: Grey value representation of DSM derived from last pulse measurement.

The laser system we are using for DSM acquisition is currently not capable to acquire multiple pulses. For this reason either the first or the last reflection of the emitted laser pulse can be measured. Since this prevents the measurement of the required data in a single pass coverage, the flight effort for laser data capture is doubled, if one aims on the acquisition of the first and last response of the emitted laser pulse. Additionally, in our examples for some areas no response could be measured at all in the last pulse mode. These regions correspond to the white areas depicted in figure 4. Besides those sensor related problems, a further differentiation of object classes like the extraction of streets or different landuse classes like grass-covered areas is not possible if only laser data is applied. For this reason in our approach the height data is integrated with multispectral imagery within a combined classification step in order to separate the required objects.

3. CLASSIFICATION OF URBAN AREAS

For the test site Color-Infrared (CIR) aerial images were available, which were taken at a scale of 1 : 5000 with a normal angle aerial camera at a focal length of 305 mm. For digitization the images were scanned at a resolution of 60 µm, resulting in three digital images in the spectral bands near infrared, red and green at a pixel footprint of 30 cm. Direct digital image acquisition would of course provide better and - even more important – reproduceable radiometric measurements, which are better suited for classification purposes. Additionally, the number of spectral bands is not restricted as if analog images are scanned for digitization. By digital cameras data can be captured in the spectral bands red, green and blue, which is necessary for lifelike visualizations as well as in the near infrared spectral range, which is important for classification. For these reasons our work is also motivated by emerging airborne systems for digital image acquisition, which will further improve our classification results.
The basic idea of the proposed algorithm is to simultaneously use geometric and radiometric information by applying a pixel-based classification. Within this classification the normalized DSM is used as an additional channel in combination with the three spectral bands. For integration of different data types, the first problem to be solved is the registration of the data sets. In order to transform the data a common system a colored ortho image is generated from the original CIR imagery based on the available DSM.

3.1. Combination of Multi-Source Data

A number of techniques are available while aiming on the combination of multisource data for scene labelling. A review of concepts and ideas for utilization of additional data sets in multipspectral classification procedures is for example given by Hahn and Stätter (1998). Generally, for our application two approaches are feasible. The available multispectral and geometric information can be combined applying the *additional channel* concept and the *hierarchical classification* approach.
Within the hierarchical classification the different types of data are applied in order to successively divide the working area into more detailed object classes (Savian and Landgrebe (1991)). Figure 8 shows the subdivision of the scene into vegetation (black) and non-vegetation (white) areas. This step was performed based on the analysis of the CIR ortho image, which is depicted in Figure 7.

'Photogrammetric Week '99' D. Fritsch & R. Spiller, Eds., Wichmann Verlag, Heidelberg, 1999.

Figure 7: Ortho image of test area.

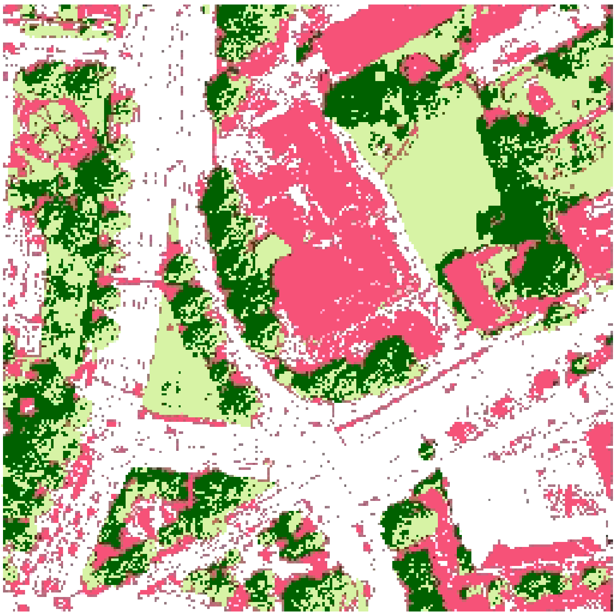
Figure 8: Vegetation and non-vegetation regions extracted from CIR aerial image.

Within the next step of the hierarchical classification these areas can be further subdivided based on the laser scanning data using the information on the local height above ground, which is provided by the normalized DSM. The vegetation regions can be separated into tree regions (high values of the normalized DSM) and other vegetation like grass-covered areas (low values of the normalized DSM). Accordingly, the non-vegetation areas can be differentiated into buildings (high areas) and non-building regions like streets (low areas). A similar approach is described by Henricson et al. (1996). They use information from color aerial images to separate elevation blobs detected in a DSM into the classes buildings and trees.
The main problem of the hierarchical or layered classification is that classification errors of the first step are propagated to the subsequent steps. Furthermore, the additional channel concept enables a more flexible processing of the data. This is the reason, why we prefer the use of this concept in our approach.

3.2. Additional Channel Concept

The main objects we are interested in are buildings, streets, trees and grass-covered areas. In order to demonstrate the insufficiencies of a standard classification, which is restricted to the analysis of multispectral information a maximum likelihood classification was applied to the CIR ortho image. The result of the separation into the required object classes is depicted in Figure 9. Figure 10 depicts the result if again the required landuse classes are generated by a standard maximum-likelihood classification. For this example the normalized DSM was introduced as an additional channel to the classification and thereby combined with the multispectral channels. Figure 9 and Figure 10 demonstrate very well that in an urban environment the classification results can be improved considerably, if the normalized DSM representing the local height above terrain is introduced as an additional data base during classification.

'Photogrammetric Week '99' D. Fritsch & R. Spiller, Eds., Wichmann Verlag, Heidelberg, 1999.



Buildings Trees Streets Grass-covered

Figure 9: Classification based on CIR imagery.

Figure 10: Classification based on CIR imagery and laser data.

3.3. Shadow Analysis

One problem while classifying areas of rough surfaces and steep slopes like they occur in urban areas is the change of spectral reflectance due to the appearance of shadowed areas. In order to avoid these problems the shadowed areas can be derived automatically based on the given DSM and used as an additional source of information during the classification process.

For the automatic detection of shadowed areas the local height, which is provided by the laser DSM as well as the elevation and azimuth of the sun at the time of image acquisition is required. The elevation and azimuth of the sun can either be determined manually by an interactive measurement of an edge of a shadowed area in the image and the corresponding object height in the normalized DSM or derived automatically from the geographical latitude and longitude of the captured area and the time of image acquisition. The result of the automatic generation of shadowed areas by applying a standard ray tracing approach is given in Figure 11.

The improvement of the classification while applying predefined shadowed areas is demonstrated in Figure 13 and Figure 14. The section of the ortho image corresponding to these examples is given in Figure 12. In Figure 13 the misclassifications in shadowed areas are clearly visible. These misclassifications can be avoided, if the class *shadow* is determined in advance based on the analysis of the given DSM and excluded from further classification like it is demonstrated in Figure 14.

'Photogrammetric Week '99' D. Fritsch & R. Spiller, Eds., Wichmann Verlag, Heidelberg, 1999.

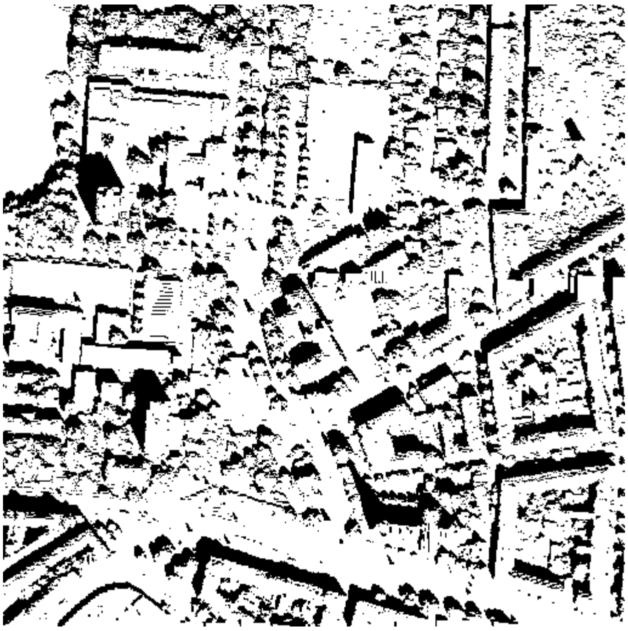


Figure 11: Shadowed areas as derived from DSM.

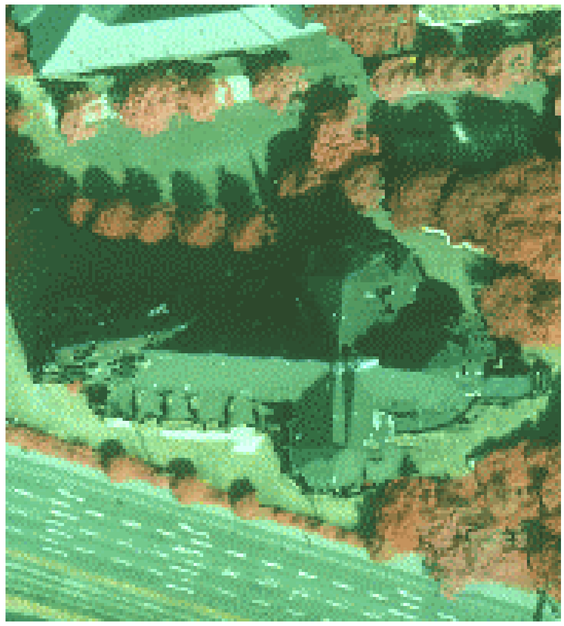
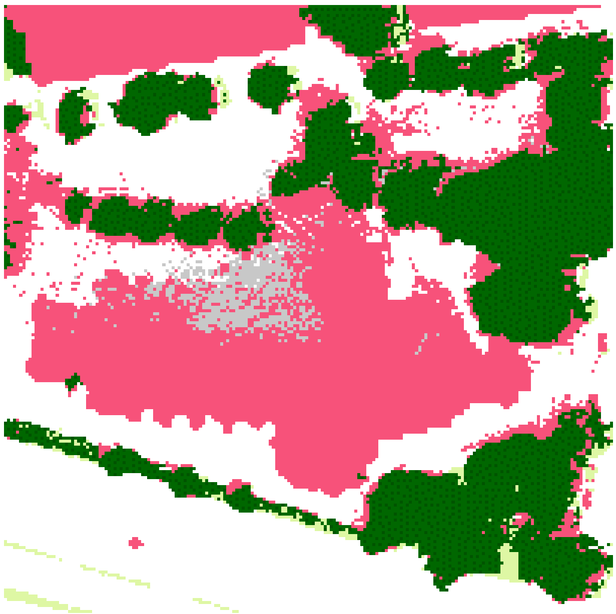
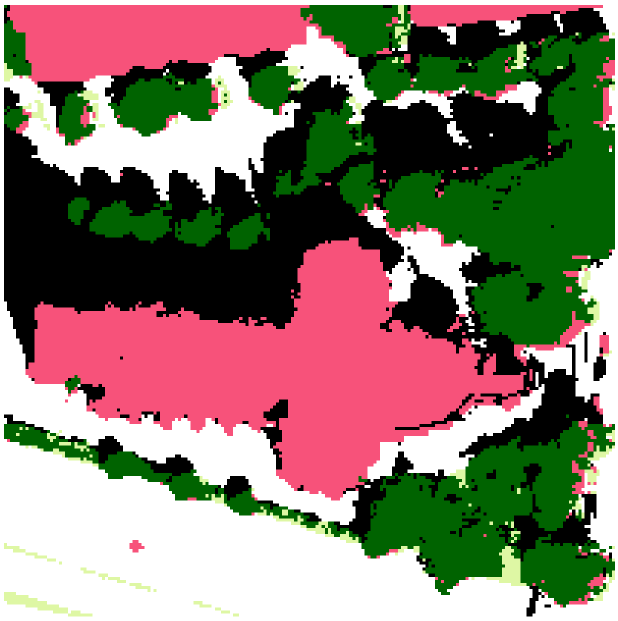


Figure 12: Example of shadowed areas.





Buildings Trees Grass-covered street Shadow Not classified

Figure 13: Classification without using shadowed areas.

Figure 14: Classification using predefined shadowed areas.

In order to avoid the *shadow* class in the final result the approach can be further refined by splitting each of the land use classes into one separate class for shadow areas and one land use class for non-shadow areas. This leads to a significant improvement of the results because the pixels in shadow areas have a completely different spectral characteristic as in non-shadow areas. After the classification the shadow and non-shadow pixels for each land use class are combined again to obtain one unique class for each type of landuse. Still, the applicability of this approach is limited due to the

'Photogrammetric Week '99'   D. Fritsch & R. Spiller, Eds.,   Wichmann Verlag, Heidelberg, 1999.

reduced radiometric quality of the input data especially at shadowed areas, which can impede the proper classification within these regions. The final result of the classification algorithm using CIR imagery and laser scanner DSM using the predefined shadow regions is given in Figure 15.

Buildings   Trees   Street   Grass-covered

Figure 15: Classification result based on CIR imagery and laser scanning using predefined shadow areas.

Figure 16: Classification result based on CIR imagery and height data from stereo matching.

Figure 17: Ortho image of test area.

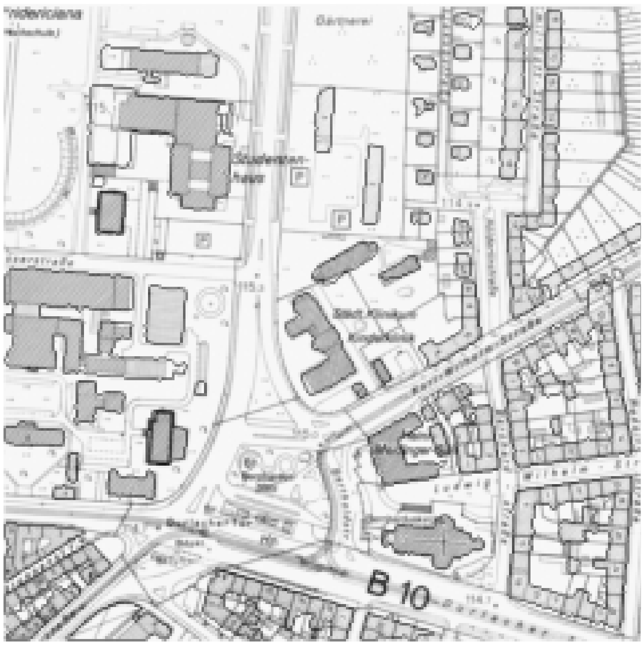


Figure 18: Comparison of existing map using classification result.

At the moment, in our framework of virtual city model generation the application of the classification based on laser data and CIR imagery is twofold, namely verification of existing GIS and initial data acquisition. Figure 18 shows an existing 1 : 5000 scale map and the buildings

classified from laser data and CIR imagery, which are represented by the black polygons. At the upper left of the area differences between the map and the result of the classification representing the actual situation are clearly visible. Even though the result of the classification might not be detailed and precise enough especially for the automatic reconstruction of buildings, the automatic detection of inconsistencies can be applied to guide an operator based revision of the GIS data. Afterwards this GIS data can e.g. be combined with the laser data for automatic 3D building reconstruction (Haala and Brenner (1998)).
In addition to the verification of an existing GIS the classification result is also used for initial data capture. Urban vegetation like single trees are usually are not available from standard GIS databases. Still they play an important role for applications aiming on visualization or simulation in urban environments. The classification result obtained by our approach can be applied to automatically determine the position of trees automatically based on the classified 'tree' regions (Brenner and Haala (1998)).

4. DISCUSSION

Within the paper the benefit of combining laser data and color aerial imagery for automatic scene labelling in an urban environment has been demonstrated. The main advantage of our approach is that the problem of dealing with the mutual complementary of the different data sources is solved implicitely by combining height and color information in a classification step. The mutual weighting of the different data is represented by the cluster centers and the distances of the feature vectors to these cluster centers. Within a supervised classification these parameters are provided for each channel and object class by the analysis of the training areas.
For all classification examples presented in the paper the training areas required in order to obtain the spectral characteristics of the different land use classes areas were digitized manually. One remaining problem in multispectral classification is that due to atmospheric effects, different spectral diffusion depending on the sunlight, different spectral characteristics of vegetation depending on season or soil, etc., new training areas have to be digitized for each data. As the digitizing of the new training areas is very time intensive we are also aiming on the automatic derivation of these areas from already existing GIS data (Walter (1998)).
Another open problem is the homogeneous quality of the different data sources, which should in principle be guaranteed over the whole test area. This premise can for example be violated for the spectral data due to occlusions near buildings within the original aerial images. At these regions no spectral information is available. Additionally, as discussed above, insufficiencies of the laser scanner can result in small areas, where the laser beam is not reflected. So even though very promising result could be achieved, especially at these regions additional effort for data acquisition and processing is required.

5. REFERENCES

Baltsavias, E., Mason, S. and D. Stallmann (1995): Use of DTMs/DSMs and Orthoimages to Support Building Extraction. In: Automatic Extraction of Man-Made Objects from Aerial and Space Images, Birkhäuser Verlag, pp. 199-210.

Brenner, C. and N. Haala (1998): Fast Production of Virtual Reality City Models. IAPRS, Vol. 32, Part 4, pp. 77-84.

Brunn A. and U. Weidner (1997): Extracting Buildings from Digital Surface Models. In: ISPRS Workshop on 3D Reconstruction and Modelling of Topographic Objects: pp. 27-34.

Haala, N. and C. Brenner (1998): 3D Urban GIS from Laser Altimeter and 2D Map Data. Int. Archives of Photogrammetry and Remote Sensing (IAPRS), Vol. 32, Part 3/1, pp. 339-346.

Hahn, M. and C. Stätter (1998): A scene labelling strategy for terrain feature extraction using multisource data. In: IAPRS, Vol.32, Part 3/1, pp. 435-441.

Henricson, O., Bignone, F., Willhuhn, W., Ade, F., Kuebler, O. et al. (1996): Project Amobe: Strategies, current status and future work. Proc. ISPRS Congress Comm. III, pp. 321-330.

Lillesand, T. and R. Kiefer (1994): Remote Sensing and Image Interpretation. John Wiley and Sons, New York.

Lohr, U. and J. Schaller (1999): High Resolution Digital Elevation Models for Various Applications. Fourth International Airborne Remote Sensing Conference and Exhibition, pp. 247-253.

Maas, H. G. (1999): The Potential of Height Texture Measures for the Segmentation of Airborne Laserscanner Data. Fourth International Airborne Remote Sensing Conference and Exhibition, pp. 154-161.

Savian, F. R. and D. Landgrebe (1991): A survey of decision tree classifier methodology. IEEE Transactions on Systems, Man and Cybernetics Vol. 21 no 3, pp. 660-674.

Schlüter, M. (1998): Multi-Image Matching in Object Space on the Basis of a General 3-D Surface Model Instead of Common 2.5-D Surface Models and its Application for Urban Scenes. IAPRS, Vol. 32, Part 4, pp. 545-552.

Walter, V. (1998): Automatic Classification of Remote Sensing Data for GIS Database Revision. In: IAPRS Vol. 32, Part 4, pp. 641-648.

Weidner, U. and W. Förstner (1995): Towards Automatic Building Extraction from High Resolution Digital Elevation Models, ISPRS Journal, 50, pp. 38-49.