

ON THE USE OF MULTISPECTRAL AND STEREO DATA FROM AIRBORNE SCANNING SYSTEMS FOR DTM GENERATION AND LANDUSE CLASSIFICATION

Norbert Haala, Dirk Stallmann and Christian Stätter
Institute for Photogrammetry (ifp)
University of Stuttgart
Geschwister-Scholl-Straße 24, 70174 Stuttgart, Germany
Ph.: +49-711-121-3383, Fax: +49-711-121-3297
e-mail: Norbert.Haala@ifp.uni-stuttgart.de

Commission IV, Working Group 2

KEY WORDS:

ABSTRACT

Airborne pushbroom systems for the direct acquisition of digital imagery enable the simultaneous capture of high quality 3D and multispectral information of a scene. The combination of geometric and radiometric information should improve tasks like DTM generation – usually solely based on stereo imagery – and landuse classification, which is traditionally restricted to the interpretation of spectral information. Within this paper these assumptions are confirmed utilizing datasets of the DPA (Digital Photogrammetric Assembly), a scanning airborne sensor which records simultaneously high spatial resolution stereo and multispectral image. Since the surface reconstruction and ortho image generation from airborne scanner imagery, which is a prerequisite for the proposed approach is different to standard procedures using full frame images, this process will be described in the first part of the paper. Afterwards the benefits of a combined evaluation of stereo and multispectral data will be demonstrated for two tasks, the generation of Digital Terrain Models and the landuse classification for thematic mapping.

1 INTRODUCTION

Procedures based on automatic stereo image matching provide a so-called Digital Surface Model (DSM), i.e. objects rising from the terrain are included in the reconstructed surface. In order to obtain a Digital Terrain Model (DTM) these objects – e.g. trees and buildings – have to be identified and removed in an automatic step. In principle the information provided by the DSM is sufficient for that purpose. Nevertheless, the high accuracy of the DSM, which is required for this purpose is very hard or even impossible to arrive at using stereo image processing. Color images indisputably contain more information than grey-value images. Hence, in the framework of DTM generation the stereo image processing should benefit from the additional use of color information. Regions like build-up or wooded areas, which are very difficult for stereo matching can be detected in advance by classifying multispectral images. On the other hand landuse classification, a task which is traditionally based on the evaluation of multispectral imagery should profit from an integration of geometric information contained in stereo images.

In order to profit from both data sources, our approach combines the multispectral and stereo information. Colored ortho images are generated using the DSM and the multispectral channels to enable a simple simultaneous use of height and color information. Afterwards the radiometric information derived from the multispectral ortho images can e.g. be combined with geometric information like roughness of the reconstructed surface or inclination of the terrain in a classification step. Thus objects like trees or buildings can be detected and considered while generating the DTM. Even more important, the extraction of landuse classes, which is required for thematic mapping applications also benefits from the combination of DSM and color information.

Our work is motivated by emerging airborne scanning systems for digital image capture, which enable the simultaneous acquisition of stereo and multispectral information in a single pass coverage. In our opinion the availability of these systems will help to overcome the shortcomings of the existing standard algorithms. The approach is demonstrated using datasets of the DPA (Digital Photogrammetric Assembly), an airborne sensor permitting the simultaneous acquisition of so-called three-line images – required for georeferencing by stereo image processing – and multispectral images in the blue, green, red and near infrared spectral range.

2 DTM GENERATION FROM AIRBORNE SCANNER IMAGERY

The procedure discussed above was applied using data provided by the DPA sensor system. The development of this system started in the late eighties and was done by the Daimler-Benz Aerospace (DASA), formerly the Messerschmitt-Bölkow-Blohm (MBB) company.

The basic parameters of the optical part of the DPA system are given in table 1. The stereo module consists of a double lens, composed of three CCD lines (forward, nadir and backward view) with 6000 pixel. Two linear arrays are optically buttomed, respectively, providing wide-angle geometry with a total width of 12000 pixel. The convergence angle is $\pm 25^\circ$ between the nadir, backward and forward looking channel, respectively. At a flying height of 2000 m above ground the ground pixel size is 25 cm. The camera is completed by a spectral module for the acquisition of multispectral images in the blue, green, red and near infrared spectral range.

The configuration of the stereo module – in combination

Stereo module	
Focal length	80 mm
Line array	2 × 6000 pix/line
Pixelsize	10 μm
Data resolution	8 bit
Convergence angle	±25°
Spectral range	515 – 780 nm
Spectral module	
Focal length	40 mm
Line array	6000 pix/line
Pixelsize	10 μm
Data resolution	8 bit
Spectral range	440 – 525 nm 520 – 600 nm 610 – 685 nm 770 – 890 nm

Table 1: Basic DPA camera parameters

with a GPS/INS module – enables the reconstruction of the flight path, i.e. the computation of the exterior orientation for each scan-line. In our system an integrated approach is applied for the georeferencing of three-line imagery by combining GPS, INS and three-line geometry. The general idea is to perform an aerotriangulation using coplanarity and collinearity condition provided by tie and control points in order to correct remaining errors of the GPS/INS component and to calibrate the whole system (Haala et al., 1998).

Additionally, the along-track stereo capability of the camera can be utilized to acquire the geometry of the sensed terrain surface. Similar to the DSM acquisition from full frame imagery, where two aerial photographs are mutually oriented to achieve stereoscopy by a relative orientation, the raw image strips acquired by the forward, nadir and backward looking channel have to be transformed for the same purpose. Each scan-line is projected to a planar surface defined by a horizontal plane at the mean terrain height using the exterior orientation parameters determined in the previous step. Within this so-called rectification step also distortions of the raw images resulting from the high frequency motion of the aircraft are eliminated. Thus the automatic measurement of parallaxes by stereo image matching is enabled. Unlike the perspective transformation of full frame images, there is no direct link between terrain coordinates and the rectified scanner imagery. For this reason corresponding points determined by image matching have to be transformed first into the original, raw scanner images. Afterwards terrain points can be calculated by spatial intersection in order to derive the required grid.

Due to the flight movement scanner images show distortions mainly caused by attitude variations of the camera. These distortions not only affect the visual impression, they also prevent the stereo viewing and result in problems for automatic image matching procedures. At least a number of points can be determined also in the original images by image matching but a dense parallax measurement, which is required for DSM acquisition is not feasible. In order to eliminate these distortions a rectification process is applied.

2.1 Rectification

For the rectification a horizontal plane located at a mean terrain height is defined (Z-plane). As indicated in figure 5 the direct method of image rectification is applied. Each

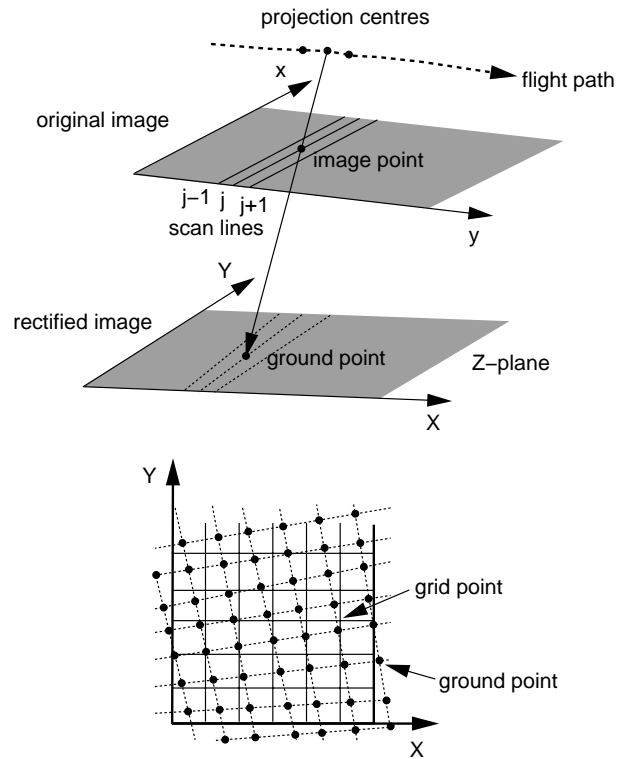


Figure 5: Image rectification

pixel in the image space is projected on the Z-plane utilizing orientation and position of the respective scan line. The resulting ground points are irregularly distributed in this plane and carry the grey value of the corresponding image points. The transfer to the regular grid is solved by interpolation using weighted averaging of grey values of all neighboring points within a certain radius. The weight is chosen reciprocal to the distance between the grid points and its neighbors.

The rectification of the stereo channels results in nearly epipolar images. Depending on the terrain height corresponding image points are shifted parallel to the flight direction. Due to a horizontal aircraft motion perpendicular to the ideal straight line of flight, a small base perpendicular to the flight direction can occur. Since this base (e.g. 10 to 20 m) is rather small compared to the base in the direction of flight, the effect for stereo viewing and image matching is negligible. Nevertheless, for airborne scanner images only quasi-epipolar images can be generated, remaining small vertical parallaxes must be expected during stereo processing.

An example for the generation of epipolar imagery from three-line pushbroom imagery is given in figures 1 to 4. Figures 1 and 2 show image strips acquired by the DPA stereo module at a flying height above ground of 2000 m. The figures present data from the forward and the nadir channel, respectively. Stereo viewing is prevented by the distortion resulting from the motion of the camera during the flight. Figure 3 and 4 show the quasi-epipolar images generated by the rectification process. These images are used for parallax measurement in the DSM generation process.



Figure 1: Original image of the forward channel

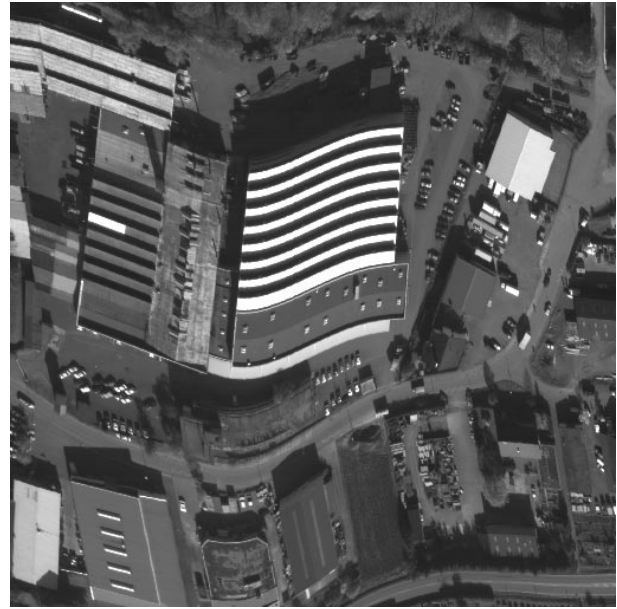


Figure 2: Original image of the nadir channel



Figure 3: Rectified image of the forward channel

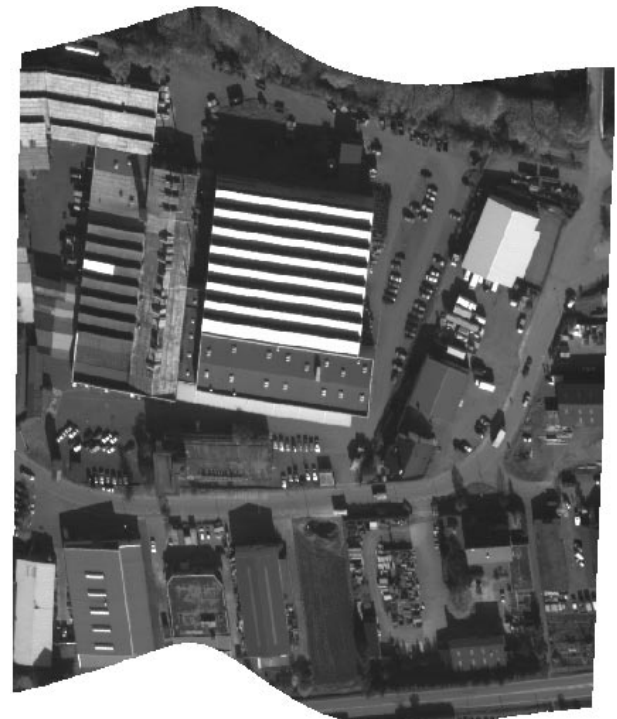


Figure 4: Rectified image of the nadir channel

2.2 Point determination in rectified images

Within rectified imagery, the application of automatic image matching as well as interactive measurement is possible with the same performance and quality than for full frame imagery. Nevertheless, in contrast to the processing of full frame imagery, the rectified images can not be directly used for point determination by spatial intersection. Within the original scanner imagery the y -coordinate of an image point coincides with the index of the scan line, i.e. the corresponding parameters of the exterior orientation can be directly accessed. After the rectification step this simple relation between pixel coordinates and corresponding exterior orientation is lost. Due to the flight movement, scan lines

projected to the Z -plane are not parallel any more. Therefore the direct correspondence between the coordinate in the direction of flight and the time of pixel acquisition is not valid any more.

Corresponding points P_1 and P_2 measured in the rectified (quasi-epipolar) images refer to one point P in 3D object space (Figure 6). Due to the assumption of an average terrain height these points have incorrect coordinates in object space: different planimetric coordinates $P_1(X_1, Y_1)$ and $P_2(X_2, Y_2)$ and the height of the rectification plane $Z = Z_1 = Z_2$. The correct position in 3D object space can be derived by a two step procedure.

1. Back projection of the corresponding points P_1 and P_2

into the original image space P' and P'' by collinearity condition.

2. Spatial intersection of the corresponding rays using of the exterior orientation results in correct object coordinates X_P, Y_P, Z_P .

Compared to the processing of full frame imagery the back projection is more expendable since the scan line in which a object point is imaged has to be determined. The location of the corresponding scan line can be found with an iterative search algorithm. After starting with an initially guessed line in the image strip, the corresponding scan line is found by minimizing the perpendicular distance to the corresponding CCD line in the camera frame.

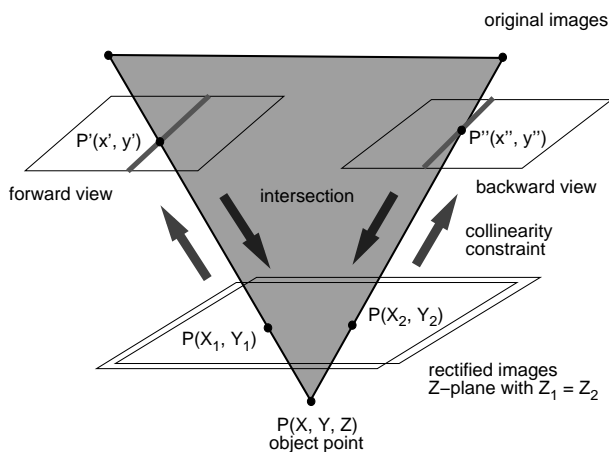


Figure 6: Point determination in rectified imagery.

2.3 Generation of ortho images

Even for large photogrammetric blocks consisting of conventional full frame imagery the corresponding image for each terrain point can be easily determined after aerial triangulation. Afterwards the terrain point can be projected into the image in order to define the corresponding grey value by applying the well known collinearity equation. This procedure is also known as indirect method of ortho image generation. In principle, the same method can be applied for scanner imagery. In contrast to full frame images a pushbroom image strip consists of a large number of single image lines, which have a width of one pixel. For this reason, the number of image lines, i.e. single images is considerably larger compared to a block consisting of full frame images. Therefore the effort for the determination of correspondences between grey values, i.e. image coordinates and terrain points during ortho image computation is much larger for pushbroom imagery. In principle, the problem of ortho image generation from pushbroom imagery by the indirect method is very similar to the problem of back-projecting points from rectified imagery to the original image strips during point determination as described in section 2.2. For computational reasons it can be better, to apply a ortho image generation by the direct method, similar to the the rectification step described earlier. For that purpose the image ray is no longer intersected with a horizontal Z -plane, but with the DSM surface.

A result of the DSM and ortho image generation from DPA imagery is presented in figure 7. The figure shows a 3D visualization of a DSM acquired by automatic intensity based

image matching between the rectified forward and backward looking channel. The overlaid ortho image was generated using the nadir looking channel. The DSM represents a region of $2 \times 3 \text{ km}^2$ and was not height exaggerated for visualization.

In the following section the combination of this geometric data with multispectral information for landuse classification and DTM generation will be discussed.

3 CLASSIFICATION BY COMBINING MULTISPECTRAL AND STEREO IMAGERY

Remote sensing for landuse classification usually deals with multispectral data sets. Up to now auxiliary information on the surface topography has mainly been used for the geometric correction of the spectral data in the framework of spectral ortho image generation. This georeferencing process of the spectral data is e.g. required if the images have to be combined additionally with site specific information provided by a GIS. Especially in hilly or mountainous terrain a radiometric correction of the multispectral imagery based on the surface orientation can be necessary, since spectral reflectance values are also influenced by the direction of illumination. For that purpose existing height data has also been used.

The basic goal of our approach is the combination of height and multispectral data for classification purposes. Spectral information is widely used as a data source for thematic mapping applications. Still the potential of spectral data is limited for these applications with respect to the accuracy and reliability of the results as well as the possibility to discriminate a large number of object categories. This was our motivation to improve classification procedures solely based on spectral data by the additional use of geometric information. Since this information on the geometry of the visible surface can be provided by stereo image matching, the availability of sensors like the DPA, which are capable to simultaneously acquire stereo and multispectral imagery will further stimulate the use of such an approach.

As described in section 2 stereo image data can be used to generate a so-called Digital Surface Model (DSM) which represents the terrain surface including all objects rising from the terrain like buildings and trees. Even though DSM are an important source of information, for most applications Digital Terrain Models (DTM) are required which are restricted to the terrain surface, i.e. objects rising from the terrain have to be detected and eliminated automatically. For that purpose spectral data can be combined with the DSM. After the classification step objects rising from the terrain like trees and buildings can be labeled and masked out precisely in order to derive a DTM from the already generated DSM. Figure 8 shows the aspired flow of operation, which applies procedures for classification and DEM generation purposes using both data types. This scheme is not yet completely implemented but several procedures are realized to enable first investigations.

For the investigations presented in this paper standard classification tools are used by introducing height information as an additional channel into the classification. A DSM usually represents an absolute height value above sea level. This type of information is only relevant for very small scale applications, e.g. the height can be used to provide a priori probabilities for certain landuse classes, since they can



Figure 7: 3D perspective view of the ortho image draped over the DSM generated from DPA imagery

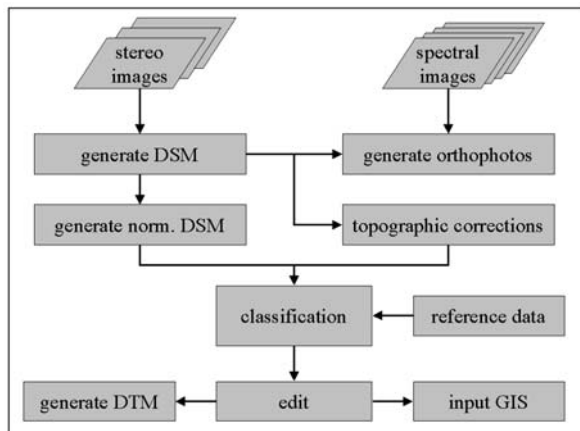


Figure 8: Workflow for applying stereo and multispectral data for DTM generation and landuse classification

be different for areas at sea level or mountainous regions. We are interested in recognizing object classes like buildings, roads, forest and agricultural areas for large scale applications. Hence information on the height above terrain level is more useful. Forest and agricultural areas as well as buildings and streets frequently have similar reflectance properties. Therefore their discrimination can be difficult, if only spectral information is available. On the other hand buildings, forest areas and single trees rise from the terrain surface; agricultural areas, water and roads are objects at terrain level. For this reason the height above terrain as an additional source of information improves the discrimination and avoids misclassifications especially for these types of objects. To make information on height above terrain accessible the so-called normalized DSM, i.e. the difference between DSM and DTM can be calculated. This surface consists of objects rising from the terrain approximately put on a plane. The required DTM, i.e. the topographic surface can be derived approximately from the measured DSM by mathematical morphology (Weidner and Förstner, 1995).

Alternatively, an existing data set can be used if available.

3.1 Classification method

The use of spectral data for landuse classification is well investigated and documented in literature, an overview is e.g. given by (Schowengerdt, 1997) or (Richards, 1993). The integration of multi-sensor data into classification are presented by (Hall, 1992), methods for the combination of multiple data sources are also discussed in general by (Hahn and Stätter, 1998). For first investigations we decided to demonstrate the potential of the approach using a rather simple method, which can be later replaced by more sophisticated procedures. Height information represented by the normalized DSM is simply introduced as an additional channel into the classification, i.e. it is processed as an additional spectral band. This enables the application of standard classification tools with only slight modifications.

The classification is based on the K-means algorithm, which is originally an unsupervised classification (Schowengerdt, 1997). Since sample data with known object types are assigned interactively by an operator for initialization, the algorithm is transformed to a hybrid or supervised classification. For data sampling only one seed point for each object class is digitized by the operator in order to limit the manual effort.

The classification procedure includes the subsequent steps:

1. Interactive acquisition of sample pixels for each spectral class.
2. Application of K-means algorithm to generate spectral clusters
3. Aggregated spectral clusters are split, if the the initial class center has changed more than a threshold, which is based on the standard deviation of the class.

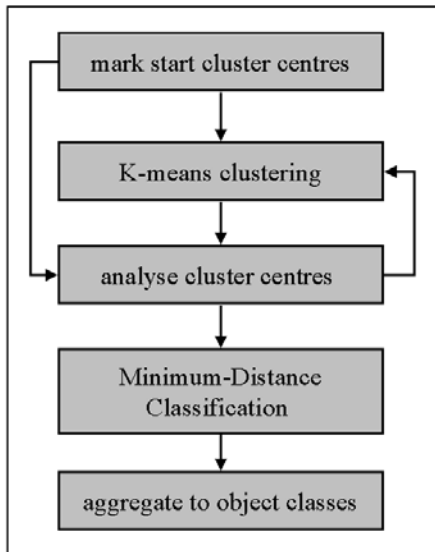


Figure 9: Workflow of implemented classification

If no cluster splitting was initiated the algorithm continues with step 4, else it is continued with step 2.

4. Classification of the whole image using Minimum-Distance Classification to all generated clusters. This step is required since for computational reasons only a certain percentage of pixels (e.g. only the pixels of each 10th row and column) are used for cluster aggregation.
5. Combination of the generated spectral classes to object classes. Due to the spectral variances of the individual object classes several spectral classes are required for each object class. Each object class is represented by several spectral classes. Hence the generated spectral classes have to be combined to object classes in the final step.

The main advantage of our algorithm is the restriction of manual interaction to the initialization step. The interpretation step, usually required for cluster aggregation is replaced by the interactive digitization of sample pixels. The object classes are determined during initialization, all classes generated during cluster splitting can be automatically re-assigned to the initial object class in the final step.

3.2 First results

The proposed algorithm was applied to data of our test site Vaihingen/Enz, which is shown in figure 7. The imagery was captured at a flying height of 2000 m resulting in a ground pixel size of 0.5 m for the multi-spectral images. Figures 10 and 11 show rectified imagery of the spectral bands red (610 - 685 nm) and near infrared (770 - 890 nm), respectively. Additionally the spectral bands blue (440 - 525 nm) and green (520 - 600 nm) were available. The normalized DSM, which was derived from the original height data by applying mathematical morphology is shown in figure 12. Even though a DSM with a grid width of 2 m like it is shown in figure 7 could be produced for our test area by applying an intensity based image matching on the DPA stereo imagery, we decided to use scanned aerial images of scale 1:13000 which were required simultaneously during that flight for DSM generation by applying the software



Figure 10: Red band



Figure 11: Near infrared band

MATCH-T (Krystek, 1991). Caused by a hardware error in the INS component of the DPA system the resolution of the gyroscopes and of the accelerometers was reduced significantly, resulting in a sensor noise approximately three times higher than the nominal value. This high frequent INS noise resulted in problems for a very dense and accurate DSM generation by feature based matching, which is more appropriate for surface reconstruction in build-up areas. The error has been eliminated in the meantime, therefore comparable results should be obtained from DPA stereo imagery in the near future.

First results of our classification using spectral and height information are shown in figure 13. During classification 82 spectral classes were generated automatically and combined to five object classes. These object classes are water (black), forest (dark grey), building (middle grey), road (light grey) and agricultural land (white). Especially for build-up areas the combination of radiometric and geometric information improves the classification compared to a standard approach, which is restricted to the spectral channels. As expected, for object classes like agricultural areas and streets, which are both at terrain level, there is no yield if the normalized DSM is introduced as additional source of information.

The result of the classification can be used to mask out the object classes forest and building during a refined DTM generation step in order to improve the approximate DTM, which was generated by morphological processing of the DSM. First results show that compared to the approximate



Figure 12: Normalized DSM

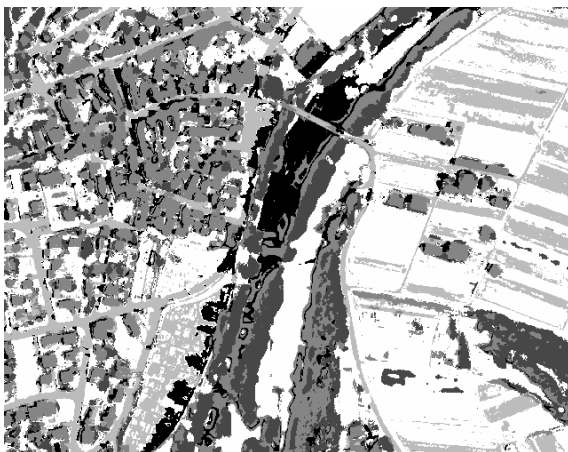


Figure 13: Classification result using spectral and height information

DTM the accuracy can especially be increased at transitions between the terrain and objects rising from that surface. In addition an unintentional smoothing of the terrain surface due to the morphologic processing can be avoided. Still further investigations are required in that area.

4 CONCLUSIONS

Within this article first results showing the potential of the combined use of stereo and multispectral data for landuse classification and DTM generation have been demonstrated. The basic idea of the proposed algorithm is to simultaneously use geometric and radiometric information by a pixel-based classification, i.e. the DSM provided by stereo image matching is used as an additional channel in combination with the spectral bands during the classification. Alternatively to the additional channel concept approaches like stratification will have to be examined. In our opinion the current trend to combine different data sources for tasks like object recognition or landuse classification will be further increased by currently emerging systems, which are able to capture different data like stereo and multispectral imagery at the same time.

REFERENCES

- Haala, N., Stallmann, D. and Cramer, M., 1998. Calibration of directly measured position and attitude by aerotriangulation of three-line airborne imagery. In: Proc. ISPRS Congress Comm. III, Columbus, Ohio, pp. 23–30.
- Hahn, M. and Stätter, C., 1998. A scene labeling strategy for terrain feature extraction using multisource data. In: Proc. ISPRS Congress Comm. III, Columbus, Ohio, pp. 435–441.
- Hall, D. L., 1992. *Mathematical Techniques in Multisensor Data Fusion*. Artech House, Inc., Boston.
- Krystek, P., 1991. Fully automatic measurement of digital elevation models with MATCH-T. In: Proceedings of the 43th Photogrammetric Week, Schriftenreihe des Institus für Photogrammetrie der Universität Stuttgart, Heft 15, pp. 203–214.
- Richards, J. A., 1993. *Remote Sensing Digital Image Analysis*. Springer-Verlag, Berlin Heidelberg.
- Schowengerdt, R. A., 1997. *Remote Sensing Models and Methods for Image Processing*. Academic Press, San Diego.
- Weidner, U. and Förstner, W., 1995. Towards automatic building extraction from high resolution digital elevation models. *ISPRS Journal* 50(4), pp. 38–49.