Orthorectification as Preliminary Step for the Fusion of Data from Active and Passive Sensor Systems

Antje Thiele, Karsten Schulz, Ulrich Thoennessen and Erich Cadario

Abstract—One of the main future tasks in the field of reconnaissance and surveillance is the data fusion of heterogeneous systems. For this purpose sensor data of airborne and spaceborne systems, which acquire the data from the same scene, with different spatial and spectral coverage are used. The use of different sensors enables the exploitation and the extraction of information on a higher level, supporting a value adding to the reconnaissance chain, which is especially helpful for the interpretation in urban terrain.

A combination of active (SAR, LIDAR) and passive (VIS, IR) sensors is considered. Because of the different imaging properties of the sensor systems it is necessary to transform the datasets into a common coordinate system.

For each imaging device a specific transformation equation system is applicable depending on the sensor parameters and supplementary metadata. In this proposal for the used devices the imaging transformation is presented, which allows a common exploitation of the sensor data for the same scenery. An orthorectification method especially for SAR images was developed and applied to airborne SAR data. The spatial accuracy of the transformed sensor data in the common coordinate system is assessed using cadastral data set.

I. INTRODUCTION

THE employment of different sensors opens the opportunity for an exploitation of complementary as well as supplementary sensor information for many purposes. By this a value added information concerning the reliability of the exploitation can be provided. In general, topographic mapping of urban areas is based on sensor data acquired from airborne platforms in nadir view under good weather conditions, e.g. aerial imagery in the Visible (VIS) and Infrared (IR) spectral channel. An alternative part of the frequency spectrum of steadily growing importance in remote sensing is the radar domain. Synthetic Aperture Radar (SAR) has some very significant advantages like independence of daytime and the large signal wavelength provides almost insensitivity to weather conditions.

Combining the data of available imaging sensors can be very complex depending on platform type (e.g. spaceborne or airborne), operation mode (active, passive), spectral range (VIS, IR, RADAR) and the used imaging device. Additionally often multi-temporal and multi-aspect data sets have to be combined. Depending on the sensor the metadata can vary in quality and can be incomplete. The preliminaries for fusion and a combined exploitation of such inhomogeneous data sets are discussed in this proposal. The proposed solution is the true orthorectification of the images.

In section II the considered sensor systems are described and the related datasets are introduced. Necessary transformation steps into a common coordinate system are discussed and shown in section III. The emphasis of this paper is to put out this very sensor specific process. In particular in the case of sensors with a different mapping behaviour adapted transformations have to be used. The assessment of results and the conclusions are presented in section IV and V.

II. CONSIDERED SENSOR SYSTEMS

Imaging systems can operate in an active or a passive mode and can be implemented in a spaceborne or airborne sensor platform. In the case of active sensors the system itself transmits the radiation, which is reflected by the investigated surface and received by the same sensor (mono static). The typical wavelengths of these sensors are in visible, near infrared and radar spectral range. Contrary to active sensors the passive sensors are detecting naturally occurring radiation, caused by the temperature of the emitting surface itself (mainly thermal infrared) or by reflecting radiation of the sun (mainly visible and near infrared). The investigated scenery covering the city of Karlsruhe (Germany) is shown in Fig. 1.

Fig. 1. Panchromatic QuickBird image of city Karlsruhe

A. Data Sets of Active Sensors

SAR sensors provide a two-dimensional mapping of the scene. However, the SAR principle requires an oblique and
side-looking viewing direction (Fig. 2). Consequently, occlusions and multi-bounce signal propagation occur frequently in urban areas [10]. Additionally, layover inevitably takes place at building locations.

The test data were acquired with the airborne AER-II sensor of FGAN [1]. The ground resolution is approximately 1 m, off nadir angle \( \theta \) about 55°, sensor altitude of about 3000 m and the distance to first range bin in slant range is 4650 m.

The employed Laser Scanner system Toposys [2] uses a pulsed laser beam and measures the time between emitting a laser pulse and the detection of the reflected echo (Fig. 3). If the beam direction and position of the sensor are known, the coordinates and the height of the reflecting object can be calculated. The scanner system consists of four major elements which are the pulsed laser, the detection device, the receiver and the time measurement unit.

The laser operates at a repetition frequency of about 83 kHz and at an eye safe wavelength of 1.56 \( \mu \)m. The fibre scanner system has a viewing angle of \( \sim 14^\circ \) and a resolution of 0.114°. The diameter of the footprint of the pulse on the ground is for example 30 cm at 300 m altitude. The position accuracy is \( \leq 0.25 \) m and the elevation accuracy \( \leq 0.15 \) m absolute and \( \leq 0.1 \) m relative. The delivered pre-processed data have a ground resolution of about 1 m.

**B. Data Sets of Passive Sensors**

A panchromatic image of QuickBird satellite was investigated as test data set (Ortho Ready Standard Imagery) with a ground resolution of about 0.6 m. The satellite has an orbit altitude of 450 km, the speed is about 7.1 km/s and the off-nadir viewing angle spans up to \( \sim 25^\circ \) [3]. The data are obtained from a linear array of CCD elements (Fig. 4). The light rays from the scene pass through the centre of the objective before reaching the sensor elements.

The considered infrared data were recorded by a frame scanner (Barr & Stroud IR-18). This mechanism has two optical scanning elements operating at right angles to one another in combination with the main imaging lens. The first element (frame-scanning element) scans the ground line-by-line and contrary to that the second element (line-scanning element) scans each line pixel-by-pixel [4]. The resulting image points lie equidistant from the perspective centre of a spherical surface (Fig. 5). The captured long wave infrared or thermal infrared signal occupies the range from 8 to 14 \( \mu \)m. The recording and display of the received monochromatic images were implemented by using analogue video signal and technology. The image data have an effective size of 768 x 500 pixel and were recorded in approximately 800 m height.

**III. TRANSFORMATION INTO A COMMON COORDINATE SYSTEM**

The data acquisition with different sensor systems opens the opportunity to accomplish a combined analysis for multiple purposes. But all of the considered sensor systems have different imaging properties, which makes it necessary to transform the data sets into a common coordinate system. This kind of processes tries to eliminate the geometric distortions of an image to get an orthographic projection. In the following different transformation processes, specified by projection geometry and sensor type, are described.
A. Orthorectification of Radar Data

The geometric characteristic of imaging radar is the projection of intensity values in slant range geometry. According to this the distance between sensor and scene objects given by signal propagation delay is defining the geometric position of incoming signals in the image. This physical principle of the sensor is responsible for specific phenomena especially in the vicinity of buildings such as foreshortening, layover and shadow [5].

Due to this kind of projection elevated objects are displaced towards the sensor (Fig. 6). This displacement depends on object height and off nadir angle. The geolocation of every pixel in the SAR-image can be determined based on the carrier navigation data and the known distance between sensor and corresponding ground point. Hence for a correct georeferencing of elevated objects, a Digital Surface Model (DSM) including buildings is needed.

![Fig. 6. Orthorectification (SAR) without (left) and with (right) elevated man made objects to be considered](image)

This is illustrated in Fig. 6 for a building point Q. If the DSM contains no elevated man made objects (Fig. 6 left) the point Q is transformed to position Q₀ in the orthoimage. Otherwise the object point is projected to the correct geolocation Q₀ (Fig. 6 right).

B. Orthorectification of Visible Data

Due to the mapping geometry of the scanning system working according to the central projection properties, elevated objects in the scenario show a misalignment away from the sensor. This effect can be observed in range images of SAR systems in an opposite way as described before. The effect for optical images is influenced by object height and distance to the sensor NADIR axis.

The necessary transformation (orthoprojection) depends on mapping scale and can be described in three separate steps [6]. Concerning the small scale scenarios in this application the earth curvature and the distortions of the national coordination systems have to be taken into account.

In a first step the pixel matrix of the orthophoto is initialized by a rastered Digital Terrain Model (DTM) or DSM. In a second step the transformation of the terrain point in a tangential Cartesian coordinate system is performed. This has its zero point in the centre of the scene on the earth ellipsoid. In a third step for each point the geometric transformation between the tangential and image coordinate system is calculated, knowing interior and exterior orientation of the sensors at the time of data acquisition. The last step includes the transformation of the grey value from the image coordinate system to the corresponding coordinate in the orthophoto matrix. The orthophoto resulting from this calculation steps directly depends on quality of the sensor model and resolution of the height model.

The image data set with known sensor model contain the parameters of the interior and exterior orientation of the satellite sensor system as well as the radiometric calibration data for each pixel element. If interior and exterior orientation for the sensor data set is not given, the user is often provided by so called Rational Polynomial Coefficients (RPCs). These polynomials are defined by 80 coefficients, which refer to the position of the sensor during data acquisition [12].

If the accuracy of the orientation data does not match the requirements an improvement of the orientation data has to be performed. Therefore measurement of Ground Control Points (GCPs) takes place. The required number of GCPs depends on the chosen refinement approach and the GCPs should be distributed equally over the image.

For an image data set with RPC orientation model such as the considered one, the improvement process can be accomplished by a polynomial approach. Apart from the influence of orientation data the result of orthorectification also depends on the quality of the height model.

![Fig. 7. Orthorectification (VIS) without (left) and with (right) elevated man made objects to be considered](image)

In Fig. 7 the building point Q is mapped to position Q’ in the image coordinate system. If the height model contains no elevated man made objects (Fig. 7 left) the point Q is transformed to position Q₀. Otherwise the object point is projected to the correct geolocation Q₀ (Fig. 7 right).

The result of orthorectification with a DSM containing buildings is also called true orthophoto. One of the problems of the orthorectification process is caused by occluded (dead view) areas in the scene. This area is marked in Fig. 7 as shadowed region and has to be replaced by information from other images or must be marked as unreliable area in the orthophoto. In some cases these areas are filled by double assignment.

C. Transformation of Infrared Data

The special geometric properties concerning the interior orientation of the imaging device (section II.B) have to be
specified by an appropriate mathematical model. The implementation of a correct mathematical model was not part of the investigation, but details can be found in [7]. Instead a simplified transformation was implemented. This process includes the registration and mosaicing of the different single frames of a scene to one image. For this purpose GCPs have to be measured in a suitable reference image (e.g. DSM) to receive 3D coordinates.

IV. ASSESSMENT OF RESULTS

The investigations of orthorectification were accomplished with the introduced data sets of the scenery Karlsruhe. In the following sub sections the results in the common coordinate system are discussed by the use of a cadastral data set as ground truth reference. To perform the orthorectification process the delivered pre-processed LIDAR-DSM described in section II.A was used.

A. Influence of LIDAR Data

A laser scanner delivers 3D point measurements in Euclidian coordinate system. For airborne systems mostly the height information is stored in a raster grid of a predefined resolution. Image cells without a measurement are interpolated by considering their neighborhood. Due to the preprocessing steps the image does not represent the original 3D information anymore. The horizontal position is slightly different and some of the height values are calculated not measured. Additionally, sometimes more than one measurement for a resolution cell exists by considering first and last echo or by combining data of several measurements of the scene during a campaign.

The overlay of cadastral data with the delivered LIDAR data in Fig. 8 shows the correct geolocation of the DSM. At some buildings in the DSM it can be observed that straight edges appear as jagged lines. These distortions have an impact on the orthorectified visible image. This will lead to errors and must be taken into account, if accuracies in the domain of a pixel are relevant.

B. Orthorectification Results of Radar Data

The orthorectification process of the radar data was performed with an enhanced approach of [8]. A constant elevation value (Fig. 9 top) and the considered LIDAR data (Fig. 9 bottom) were used as height information in the computing process.

The initial delivered carrier navigation data were improved with the help of GCPs. These points were automatically calculated based on a correlation of predicted shadow and layover areas with potential shadow and layover regions in the original SAR image in slant range geometry. The prediction of these areas is based on the high resolution DSM. The result of the correlation leads to a correction of the carrier navigation data and is performed in an iterative manner. In the investigated dataset e.g. a displacement of about 20 m for the highest building was successfully corrected with this method.

As described in section II the layover effect at building walls occurs as bright area in front of the buildings (Fig. 9 top). To shift these intensity areas to the building location the DSM was used. Due to the fact that the intensity values in the layover areas result from multiple ground points, the shifting leads to information gaps in front of the buildings (black areas in Fig. 9 bottom).

C. Orthorectification Results of Visible Data

The orthorectification of the scenes was performed on the basis of a constant elevation value, with a DTED2 (Digital Terrain Elevation Data - Level 2) data set and the introduced high resolution LIDAR DSM. The results of the orthorectification with a commercial software tool (ERDAS Imagine, LPS) are depicted in Fig. 10a-d. The expected differences caused by the different height models are clearly to identify at building locations. It can be seen that if the cadastral information coincide all the more with building footprints, the height values in the used height model are closely describing the real surface. This is particularly visible in the result based on the LIDAR height model (Fig. 10 c). In this true orthophoto the existing dead view areas were filled with double assignments.

In the result based on LIDAR DSM a part of the building facade between bottom line and roof is visible. This effect can be explained by invalid height values in the DSM or by a misregistration between image and DSM. A registration error
is caused either by a geolocation error of the DSM or by an imprecise reconstruction of the mapping geometry from interior and exterior orientation. The overlay of DSM and building polygons show no gaps (Fig. 8), so it can be assumed that the observed effects are caused by deficits in the reconstruction of the mapping geometry. For the investigated QuickBird image the camera model was given by an RPC data set. By improving the RPC data set a more exact orthophoto can be computed.

Usually a refinement of the orthorectification results is performed by recalculation of orientation parameter. If the sensor orientation is given by RPCs, a refinement of these coefficients is not recommended, because of the correlation between the coefficients among themselves. In this case a polynomial approach should be chosen [9].

A comparison of the orthorectified QuickBird image without and with correction polynomial is shown in Fig. 10 c, d. It can be observed that the misalignment in vertical direction was corrected for high buildings. The region which showed part of the facade at the front of the buildings (c) was replaced by roof areas (d). In Table I the Root Mean Square Error (RMSE) for a set of control point pairs is shown for the different orthorectification results (b-d). The reference points were corners of cadastral building footprints, the points in the VIS images were measured at the corresponding roof corner of the building.

<table>
<thead>
<tr>
<th>TABLE I: RMSE BUILDING CORNERS</th>
</tr>
</thead>
<tbody>
<tr>
<td>DTED (b)</td>
</tr>
<tr>
<td>RMSE [m]</td>
</tr>
</tbody>
</table>

In the last step the degree of the correction polynomial was investigated. For this scene a correction polynomial of first and second order was calculated based on measured GCPs. In the comparison of these results partly differences were observed. The orthophoto based on first order correction polynomial delivered better results in comparison with overlaid cadastral data. This can be attributed either to the distribution of the GCPs, to inaccuracies determining the GCPs or to an over-parameterisation [9].

The choice of the correction polynomial is influenced by the precision of the mapping geometry, the variation in the height of the scene, the geometric resolution and the required accuracy.

D. Transformation Results of Infrared Data

Due to the low flight height and the limited Field of View (FOV) the projective effects are limited; therefore the simplified transformation was applied (section III).

Fig. 11 Mosaiced infrared image overlaid with cadastral information

In Fig. 11 the result of the transformation overlaid with cadastral data is shown. The movement around the roll axis of the airborne platform has strong effects to the image data. This is documented by an undulated run of building edges which are normally observable as straight lines in the image. These effects were not compensated by the applied transformation.

V. COMBINED EXPLOITATION

The main advantage of a fused exploitation of multi-sensor data is the possibility to combine complementary information for the same object. Different object features or
characteristics can be achieved only by complementary sensors. For example emission in the long wave infrared range is helpful to identify the operational condition of an object. In Fig. 12 the comparison of IR image in a pseudo colour presentation with the orthorectified QuickBird image is depicted.

Some of the flat roof buildings show clearly enhanced emission due to parts of heating and air conditioning systems located on the roof. Especially the IR channel has been investigated to analyse the urban infrastructure or the technical status of buildings. Typical examples are the examination of the operational condition of a district heating network concerning leakages or the heat loss of buildings concerning energy saving actions.

The imaging with SAR sensor is able to deliver information nearly independent from weather and illumination conditions, at night time and also high spatial resolution images can be achieved from high distances. This may be important in applications like disaster management. In time critical events SAR is one of the most suitable remote sensing techniques for gathering useful actual data under certain circumstances. The resolution of the used SAR image is relatively low, so that only large scale objects could be detected. But this is no principle restriction, because state of the art sensors offer clearly higher spatial resolutions like the airborne PAMIR sensor of FGAN [11]. In Fig. 13 for example the eaves and the ridge of the roof can clearly be identified as linear structures in the SAR data.

VI. CONCLUSIONS

Aim of the investigation was to analyse necessary steps to enable the fusion and combined exploitation of inhomogeneous data sets. In particular for the exploitation of multisensor data the orthorectification of the data with high resolution DSM data is a necessary prerequisite. In particular in the case of sensors with a different mapping behaviour it was shown that adapted transformations have to be used.

The quality of the used DSM has a not negligible influence to the orthorectification process. Especially in urban areas a high resolution DSM including all man made objects is necessary to achieve the aspired true orthoprojection. These high resolution DSMs nowadays are commercially available by LIDAR scanning systems. The orthorectification of the infrared images has to be improved by considering the exact camera model. For a precise orthorectification of all the described sensor data a refinement of the given sensor orientation is necessary, which has to be achieved by measurement of GCPs. For SAR images, an orthorectification method including automatic calculation of these points was developed.

ACKNOWLEDGMENT

We want to thank Prof. Dr. Ender and Dr. Brenner (both FGAN-FHR) for providing the SAR data acquired with the AER II experimental system [1] and the PAMIR sensor [11] of FGAN.

REFERENCES
