METHODS

Study area and data description

The data set was captured over Vaihingen in Germany. Area consists of three test areas: Area 1 (Inner City), Area 2 (High Riser) and Area 3 (Residential Area).

Preprocessing

For all three Vaihingen data sets the DTM with 0.25 m grid cell was produced from ALS using LAStools software. The orthoimages with spatial resolution of 0.25 m were created using DTM and aerial images with the SOCET SET. By calculating the difference between DSM and DTM a nDSM was also generated for all three test areas.

Automatic building extraction

The main steps of automatic building extraction are shown in Figure 1. The input data for building vectorisation were calculated nDSM and vegetation mask, obtained from orthoimages. Before final vectorisation of buildings, several building masks were produced.

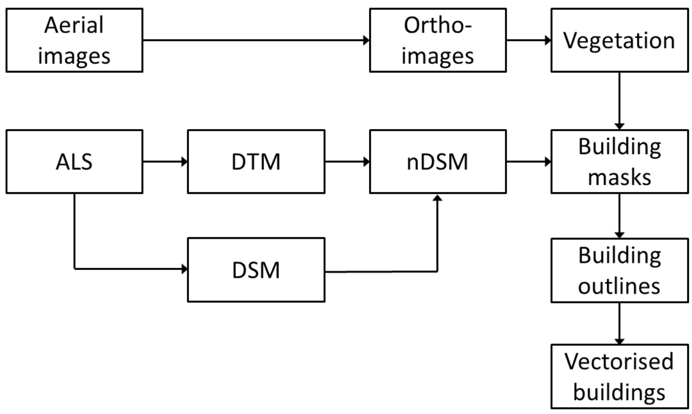

Figure 1. Flowchart of the building extraction.

By applying the 1.5 m threshold to nDSM a nDSM high objects mask was produced. Apart from buildings the mask also included other objects higher than 1.5 m. Such objects were mainly vegetation. By removing vegetation from high objects mask the initial building mask was created. Vegetation was extracted from orthoimages with the object-based methodology as described in section 2.4. Faster approach for the production of vegetation mask would be calculation of the NDVI from red (R) and infrared (IR) band of the orthoimage (1):

$$\mathrm{NDVI} = \frac{IR - R}{IR + R} \tag{1}$$

We used NDVI to remove vegetation from nDSM high objects mask only in Area 3. For this area the 0.17 NDVI threshold was applied to produce the vegetation mask.

In further processing we deal with three different masks: the nDSM high object mask, initial building mask and the mask for the building area test. Additionally we used building outlines image. The elaboration of mask for the building area test and the building outlines image is in detail discussed in Grigillo et al. (2011) within section 3.3. The mask for the building area test is made with a series of morphological operations from the initial building mask. Morphological operations were used to remove irregularities from the initial building mask. Irregularities are the consequence of different objects, which are higher than 1.5 m and do not represent buildings (fences, vehicles), noise in nDSM, errors in the calculation of the vegetation etc. The mask for the building area test contains only buildings in the test area.

For the building extraction the building outlines image was also included in the procedure. The nDSM high objects mask, the mask for the building area test and serial of morphological operations were employed to produce the nDSM building mask, where one object in the mask represented one building or several buildings, if these touched each other. Finally, the building outlines were produced from the nDSM building mask.

Since the building outlines image was in raster format, the contours were not presented with straight lines as expected for buildings. For this reason, the final building outlines were vectorised using Hough transform, separately for each object within the building outlines image. We implemented two procedures for building vectorisation: vectorisation of buildings with rectangular shape and vectorisation of non-rectangular shaped buildings. The full scenario of vectorisation procedure of rectangular buildings with Hough transform included the following steps:

- Orientation of the main building axis was determined by straight line which included the largest number of pixels that belonged to the building outline (Figure 1 (a)).
- Detection of all straight lines parallel to the main building axis that described individual building in the building outlines image. The detection of straight lines went on until at least four pixels were included within the line or until at least two parallel straight lines were detected. Straight lines were detected with the mutual distance of 0.5 m.
- Since we assumed that buildings have rectangular sides (which however is not true for some of the buildings in Area 1), in the same way also all rectangular straight lines describing building outline were detected. In this case the threshold when to stop straight line detection along the axis perpendicular to the main building axis was slightly lowered (until 3 pixels were in the line or at least two rectangular straight lines were detected). Figure 1 (b) shows parallel and perpendicular lines in the building outline.
- Intersections among rectangular straight lines were detected and out of them rectangles were made (Figure 1 (c)).
- All obtained rectangles do not necessarily correspond to buildings (Figure 1 (d) shows constructed rectangles on the mask for the building area test). To detect false positives we used the following criteria: if the ratio between the surface area obtained by the intersection of the rectangle with the mask for the building area test and the surface area calculated from the rectangle's corners was larger than 0.5, the rectangle was retained, otherwise it was discarded (Figure 1 (e)). When dealing with building interiors, the surface area test was carried out by negating the mask for the building area test.
- Vectorised building was obtained using the outline of retained rectangles. Figure 1 (f) shows the vectorised building on the DSM.

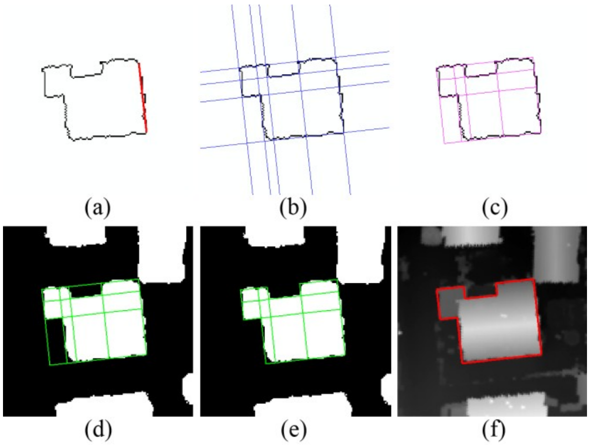

Figure 1. Rectangular building vectorisation procedure.

Rectangular building vectorisation procedure is shown in Figure 1. Figures 1 (a), (b) and (c) are shown as negatives.

As mentioned before, not all the buildings in Area 1 have rectangular shape so the procedure for extraction of that buildings was modified. The input data were again building outlines image and the mask for the building area test and the procedure operates on each separate object in building outlines image. The full scenario of vectorisation procedure of non-rectangular buildings included the following steps:

- Using Hough transform we detected all the lines that described building outlines. (Figure 2 (a and b)).
- In the next step we calculated the intersections of detected lines. For the further process of triangulation we wanted to avoid extracting lines that do not describe the building outlines so we kept only intersections that were within the distance of 1 m from the building outlines. Also the intersections that were within the distance of 1 m to one another were replaced by their average coordinates to avoid excessive number of intersections for further processing. Kept intersections are shown in Figure 2 (c) in blue colour. All the remaining intersections are shown in green colour.
- From the intersections that were kept by the process, the Delaunay triangulation was calculated. The triangles of the Delaunay triangulation are shown in Figure 2 (d).
- For each triangle two surface areas were calculated. The surface area of the intersection of the triangle with the mask for the building area test and the surface area, calculated from the triangle corners. If the ratio between calculated areas was higher than 0.6, the triangle was kept for the building vectorisation. The final triangles are shown in Figure 2 (e).
- Vectorised building was obtained by outline of all the kept triangles. Figure 2 (f) shows the vectorised building on the DSM.

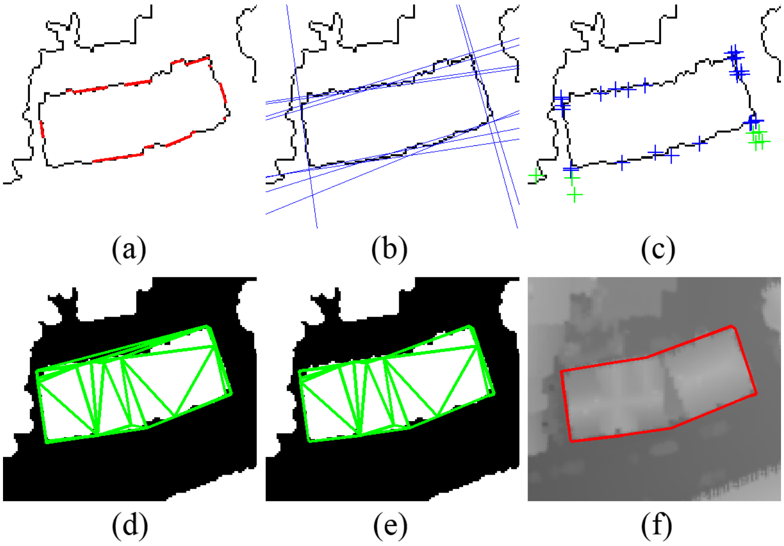

Figure 2. Non-rectangular building vectorisation procedure.

Non-rectangular building vectorisation procedure is shown in Figure 2. Figures 2 (a), (b) and (c) are shown as negatives.

Automatic vegetation extraction

Object-based classification was performed with iterative object based image analysis implemented on orthoimages with three spectral bands (IR, R, G); as the fourth band a nDSM layer was added. The resulting combined image (3+1) gave best results to locate and delineate trees and other green areas, marked as natural ground. All three test sites contain several tree species, implying large variation of distance between trees, crown sizes and surrounding elements, resulting in challenging and realistic cases for mainly deciduous tree detection.
The data has been processed with ENVI EX Feature Extraction module that includes object-based approach implemented in two steps. In the first step the segments are created based on the spectral signature of the different image parts (segmentation), while in the second step the segments are analysed and classified into the classes (classification).

Although many image segmentation methods have been developed, there is a strong need for new and more sophisticated segmentation methods to produce more accurate segmentation results for urban object identification on a fine scale (Li et al., 2011). Segmentation can be especially problematic in areas with low contrast or where different appearance does not imply different meaning. In this case the outcomes are represented as wrongly delineated image objects (Kanjir et al., 2010). Object features of type vegetation were examined at multiple segmentation scales which yielded different shapes of image-objects. After the segmentation different attributes (spatial, spectral, geometrical, texture) were calculated for each segment. NDVI (see Equation 1) was assigned to each segment.

After image was segmented segments were classified by creating a rule set according to suitable segment attributes. Texture, shape and contextual features are keys to the identification of trees in urban scenes (Ardila et al., 2012). Due to the complexity of the scene and several factors limiting vegetation detection, we detected vegetation objects using the following processing steps: 1.) classify 'visible' vegetation (i.e. vegetation within unshaded areas) and shadows, 2.) create mask of shadows, 3.) detect vegetation under shadows, 4.) merge vegetation results from visible vegetation and vegetation under shadows, 5.) divide vegetation results into trees and natural ground (high and low vegetation). Figure 3 presents a diagram of the workflow described in this section.

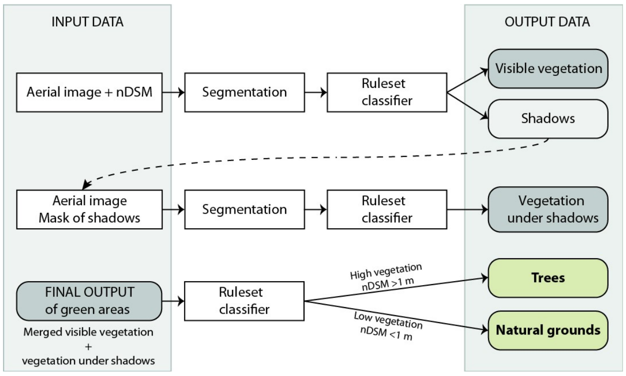

Figure 3. Workflow for urban vegetation extraction with object-based approach.

Objects of visible vegetation were detected based on NDVI which is high on vegetated areas. Simultaneously, we detected also class "shadows"; these dark areas may also contain vegetation, however in the shades the contrast between the image objects of adjacent land covers was smaller, therefore shaded vegetation areas could not be detected exclusively from NDVI. The rule set of attributes used in this procedure that correspond to class shadows is listed in Table 1. The threshold value for each attribute was identified through visual interpretation.

Class	Attribute	Value
Visible vegetation	NDVI	>0.15
Shadows	nDSM	<2.5
	R band value	<800
	area	>10
	NDVI	<0.15
Vegetation under shadows	R band value	<725
	G band value	>260
	NDVI	>0
	area	>1.5

Table . Rule set of attributes for three given classes.

Later, shadowed areas were masked out from the whole image scene and analysed in the same way as visible vegetation. Shadows produced by high objects (high buildings or high vegetation) are elemental in urban vegetation analysis. They represent a great challenge in remote sensing and symbolize a factor that considerably influences the results. Although land use under shadows has low spectral separability, vegetation under shadows still appears to have higher NDVI than other land use classes under shadows. Further promising characteristic we successfully utilize is that objects that are assigned to vegetation under shadows class have specific geometrical and spectral properties (Table 1).

The set of all vegetation objects was created by merging segments of obtained visible vegetation and vegetation under shadows. Last step was to distinguish trees from the natural grounds. It was based on nDSM value of segments – values > 1 m were assigned to trees and values < 1 m were assigned to natural grounds.

When final vegetation mask was defined we developed a program that reads this mask in intersection with nDSM data as an input for the purpose of single tree detection. First mask is processed based on watershed segmentation algorithm in order to obtain segments that represent higher vegetation objects – tree crowns. Segments of tree crowns are later merged together based on the proximity of their peaks. When these segments are defined, radius is calculated considering the barycentre of each segment. In the last stage geometrical shapes of single trees are plotted to represent positions of each tree in our study area.