ISPRS Test Project on Urban Classification and 3D Building Reconstruction

December 12, 2013

ISPRS - Commission III - Photogrammetric Computer Vision and Image Analysis

Working Group III / 4 - 3D Scene Analysis

http://www.commission3.isprs.org/wg4/

Franz Rottensteiner[1], Gunho Sohn[2], Markus Gerke[3], Jan Dirk Wegner[4]

[1] Institute of Photogrammetry and GeoInformation, Leibniz Universität Hannover,
Nienburger Str.1, 30167 Hannover, Germany;
E-mail: rottensteiner@ipi.uni-hannover.de

[2] GeoICT Lab; Earth and Space Science and Engineering Department, York University,
4700 Keele St., Toronto, M3J 1P3, Canada;
E-mail: gsohn@yorku.ca

[3] University of Twente, Faculty ITC; EOS department, Hengelosestraat 99, P.O. Box 6,
7500 AA Enschede, The Netherlands;
E-mail: gerke@itc.nl

[4] Institute of Geodesy and Photogrammetry, Swiss Federal Institute of Technology Zurich,
Wolfgang-Pauli-Strasse 15, 8093 Zurich, Switzerland;
E-mail: jan.wegner@geod.baug.ethz.ch

1. Overview

The automated extraction of urban objects from data acquired by airborne sensors has been an important topic of research in photogrammetry for at least two decades. One observation that can be made is that most of the papers dealing with urban object extraction focus on a single object class, e.g. buildings, roads or, less frequently, trees. Typically, results are published for a few test sites available to the authors. Results presented by different authors typically refer to different scenes so that the quality of the results achieved using different algorithms cannot really be compared. There have been attempts in the past to distribute data sets for benchmarking object extraction techniques, e.g. the Avenches data set of ETH Zurich or OEEPE/EuroSDR data sets for building extraction or road extraction. Even if these data sets were still available, they would be outdated due to the fact that they are based on scanned aerial images acquired by analog cameras. Today, the transition to digital aerial cameras has nearly been completed, so that there is a need for new standard test sites for urban object extraction making use of the full benefits of modern airborne data, including multiple-overlap geometry, increased radiometric and spectral resolution, and (in the case of airborne laserscanner data) the recording of multiple echoes. Two such test sites, each containing several test areas for which reference data are available, are provided for the participants in this project in order to evaluate techniques for the extraction of various urban object classes. The test data are described in detail in Section 2 of this document.

The participants in this project can choose any of the following tasks (described in detail in Section 3 of this document:

1) **Urban Object Detection:** In this context, it is the task of the participants to determine the outlines of objects in the input data. Reference data are available for a variety of object classes, including buildings, roads, trees, and cars. The participants may choose to detect single object classes, or they can try to extract several object classes simultaneously, for instance to benefit from context information, i.e. the information contained in the mutual arrangement of objects in complex urban scenes such as those distributed in this project. The reference data will be used to determine the completeness, correctness, and quality of the results, both on a per-area-level and on a per-object level.

2) **3D Building Reconstruction:** The participants shall reconstruct detailed 3D roof structures in the test areas. Detailed 3D models of roofs are available as reference data. They will be used to evaluate the quality of the roof plane segmentation process as well as the geometrical accuracy of the outline polygons of the roof planes.

The participants shall submit their results to the organizers of the test, who will compare these results to the reference data and inform the participants about the results of this comparison. First results of this project shall be presented at the XXIInd ISPRS Congress in Melbourne in 2012. The final results will be published in an international photogrammetric journal.

2. Data Information

2.1. Data Set 1: Vaihingen

2.1.1. Overview

The first test data set was captured over Vaihingen in Germany. The data set is a subset of the data used for the test of digital aerial cameras carried out by the German Association of Photogrammetry and Remote Sensing (DGPF) [Cramer, 2010]. It consists of three test areas for which reference data for various object classes are available [Spreckels et al., 2010] (yellow areas in Figure 1) and a larger test site **"Roads"** for road extraction (blue area in Figure 1):

- **Area 1: "Inner City":** This test area is situated in the centre of the city of Vaihingen. It is characterized by dense development consisting of historic buildings having rather complex shapes, but also has some trees (Figure 2a).

- **Area 2: "High Riser":** This area is characterized by a few high-rising residential buildings that are surrounded by trees (Figure 2b).

- **Area 3: "Residential Area":** This is a purely residential area with small detached houses (Figure 2c).

- **"Roads":** This area encloses all the other test sites and can be used for testing urban road extraction techniques.

For each of these test areas, the following data are distributed:

- **Digital Aerial Images and Orientation Parameters:** The images are a part of the Intergraph/ZI DMC block with 8 cm ground resolution [Cramer, 2010]. Each area is visible in multiple images from several strips. The orientation parameters are distributed together with the images.

- **Digital Surface Model (DSM) and True Orthophoto Mosaic:** The DSM was generated from the original images by dense matching using the Match-T software [Lemaire, 2008]. Based on this DSM, a true orthophoto mosaic was generated. The two data sets are defined on the same grid, having a ground resolution of 9 cm.

- **Airborne Laserscanner Data:** The test area was covered by altogether 10 strips captured with a Leica ALS50 system. Inside an individual strip the average point density is 4 pts/m^2 [Haala et al., 2010]. In addition to the original point cloud, a digital surface model (DSM) is also made available.

In the following sections, the data are described in more detail.
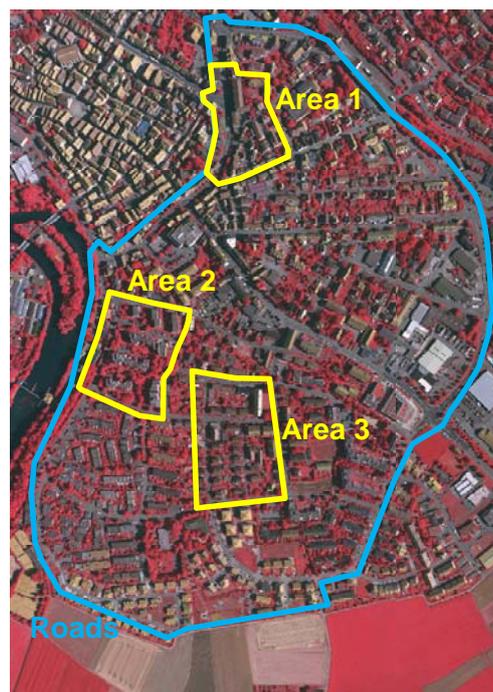
Figure 1: The Vaihingen test areas overlaid to images 10050105 and 10050107.

2.1.2. Digital Aerial Images

The digital aerial images are a part of the high-resolution DMC block of the DGPF test [Cramer, 2010]. They were acquired using an Intergraph / ZI DMC by the company RWE Power on 24 July and 6 August 2008. In total, the block consisted of five overlapping strips with two additional cross strips at both ends of the block. The test areas are visible in four of these strips, namely strips 3, 4, 5, and the cross-strip 25. Figure 3 shows the configuration of the image strips the test areas are visible in.
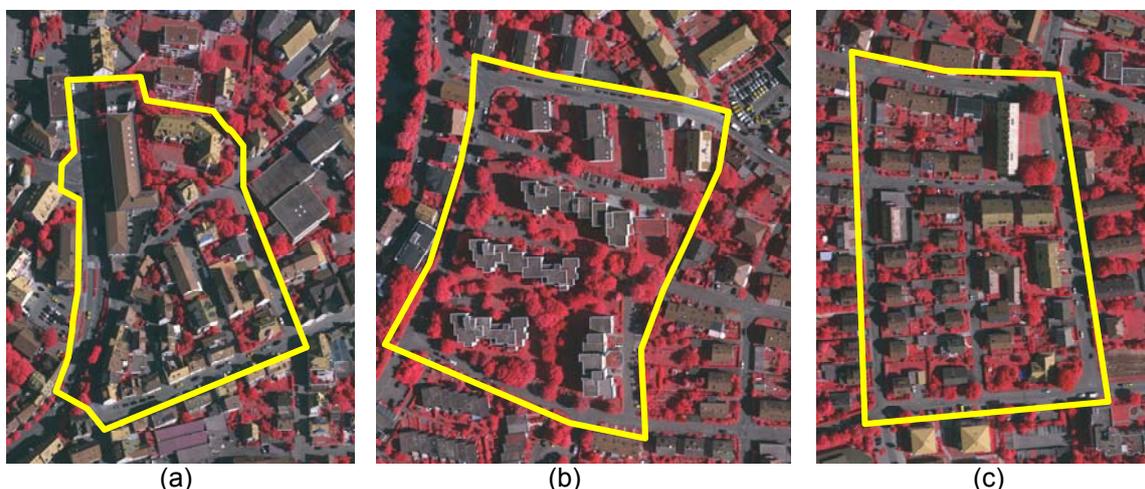
Figure 2: The three test sites in Vaihingen: a) Area 1, (b) Area 2, (c) Area 3.

Table 1 shows the flight parameters of the block, whereas whereas Table 1 gives an overview about the images the test areas are visible in. The images are pan-sharpened colour infrared images with a ground sampling distance of 8 cm and a radiometric resolution of 11 bits. They are provided as 16-bit RGB Tiff files.

Camera	Focal length	Flying height above Ground	Forward overlap	Side lap	GSD	Spectral bands	Radiometric resolution
DMC	120 mm	900 m	60 %	60 %	8 cm	IR – R - G	11 bit

Table 1: Flight parameters of the Vaihingen 8 cm DMC block. GSD: Ground Sampling Distance.

	Strip 3	Strip 4	Strip 5	Strip 25
Area 1	10030060*10030061 10030062 10030063*	10040082* 10040083 10040084 10040085*	10050104*10050105 10050106	10250130*10250131 10250132 10250133*
Area 2	-	10040081* 10040082 10040083 10040084*	10050103*10050104 10050105 10050106*	-
Area 3	-	10040082* 10040083 10040084	10050104*10050105 10050106	10250132*10250133 10250134 10250135*
Roads	10030060* - 10030063*	10040081* - 10040085*	10050103* - 10050107*	10250130* - 10250135*

Table 2: Overview about the images the Vaihingen Block. The asterisk (*) means that the area is only partially visible in that image.
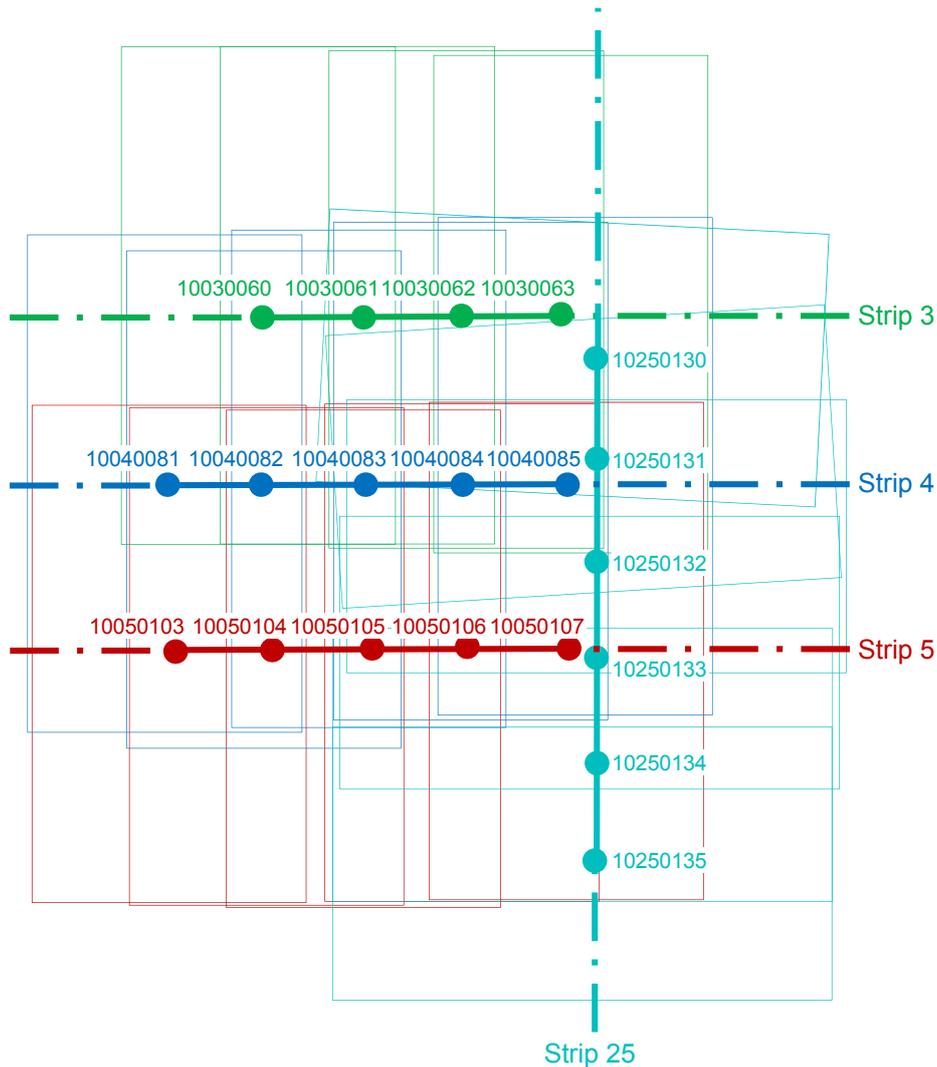
Figure 3: Image configuration for the Vaihingen test site. **Orientation Parameters**

Two coordinate systems are defined in the focal plane (Figure 4):

1) **File Coordinate System (*row, col*):** This coordinate system is defined in the image file. Its coordinate axes are parallel to the row and column directions of the digital image. Its units are [pixels], and its origin is at the left upper corner of the left upper pixel. Thus, the centre of the upper left pixel has the file coordinates $(row, col)^T = (0.5, 0.5)^T$. The Principal Point **PP** has the file coordinates $(row_{PP}, col_{PP})^T$.

2) **Camera Coordinate System (x_c, y_c):** This is a mathematically positive system. Its x_c-axis is parallel to the *col*-axis of the file coordinate system, whereas its y_c-axis is parallel to the *row*-axis of the file coordinate system, but points into the other direction. Its units are [mm]. The centre of the camera coordinate system is the principal point **PP**, which thus has the camera coordinates $(x_{PP}, y_{PP})^T = (0.000, 0.000)^T$.

The projection centre is situated on a straight line orthogonal to the focal plane and passing through the principal point; its distance from the focal plane is the focal length *f*. The file and camera coordinate systems are related via Equations 1 and 2:

$$x_c = \Delta \cdot \left(col - col_{PP} \right)$$
$$y_c = -\Delta \cdot \left(row - row_{PP} \right)$$
(1)

$$col = col_{PP} + \frac{x_c}{\Delta}$$

$$row = row_{PP} - \frac{y_c}{\Delta}$$
(2)

In Equations 1 and 2, Δ is the pixel size of the camera in [mm]. The coordinates of the principal point in the file and the camera coordinate systems, the pixel size Δ, and the focal length *f* for the images in the test block, defining the **interior orientation** of the images, are shown in Table 3.

The **object coordinate system** (*X, Y, Z*) is the system of the Land Survey of the German federal state of Baden Württemberg, based on a Transverse Mercator projection. The **exterior orientation** of the images is given by the object coordinates of the projection centres $\mathbf{P}_0 = (X_0, Y_0, Z_0)^T$ and three rotational angles (ω, φ, κ), where ω is the primary rotation about the *X*-axis,

Figure 4: File and camera coordinate systems.

φ is the secondary rotation about the rotated *Y*-axis, and κ is the tertiary rotation about the rotated *Z*-axis. The three rotational angles (ω, φ, κ) are related to the rotational matrix **R** according to Equation 3:

$$\mathbf{R} = \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix} = \begin{pmatrix} \cos\varphi \cdot \cos\kappa & -\cos\varphi \cdot \sin\kappa & \sin\varphi \\ \sin\omega \cdot \sin\varphi \cdot \cos\kappa + \cos\omega \cdot \sin\kappa & -\sin\omega \cdot \sin\varphi \cdot \sin\kappa + \cos\omega \cdot \cos\kappa & -\sin\omega \cdot \cos\varphi \\ -\cos\omega \cdot \sin\varphi \cdot \cos\kappa + \sin\omega \cdot \sin\kappa & \cos\omega \cdot \sin\varphi \cdot \sin\kappa + \sin\omega \cdot \cos\kappa & \cos\omega \cdot \cos\varphi \end{pmatrix}$$
(3)

R is defined to rotate from the camera coordinate system to the object coordinate system. Using **R**, the projection centres $\mathbf{P}_0 = (X_0, Y_0, Z_0)^T$, and the focal length *f*, the relation between the camera coordinates (x_c, y_c) and the object coordinates $\mathbf{P} = (X, Y, Z)^T$ of a point is given by Equations 4 and 5:

$$\mathbf{P}_C = \left(X_C, Y_C, Z_C \right)^T = \mathbf{R}^T \cdot \left(\mathbf{P} - \mathbf{P}_0 \right)$$
(4)

$$x_c = -f \cdot \frac{X_c}{Z_c}$$

$$y_c = -f \cdot \frac{Y_c}{Z_c}$$
(5)

Camera	file coordinate system			camera coordinate system			pixel size Δ [mm]
	row_{PP} [pixel]	col_{PP} [pixel]	*f* [pixel]	x_{PP} [mm]	y_{PP} [mm]	*f* [mm]	
Intergraph/ZI DMC	6912.0	3840.0	10000.0	0.000	0.000	120.000	0.012

Table 3: Interior orientation of the digital images of the Vaihingen Block.

The parameters of the exterior orientation of the images are shown in Table 4. These parameters can also be found on the file `daporo.dat` in the same directory as the images, where each line corresponds to an image in the format:

```
image   omega [gon]   phi [gon]   kappa [gon]   X0 [m]   Y0 [m]   Z0 [m]   f [mm]
```

The file `daporp.dat` in the same directory as the images also contains the exterior orientation parameters, but it contains the rotational matrix **R** rather than the rotational angles. Each image corresponds to a group of three lines in the format:

```
image     f [mm]     X0 [m]      Y0 [m]      Z0 [m]

r11       r21        r31         r12         r22

r32       r13        r23         r33
```

For the purpose of this test, the camera can be assumed to be without any distortion. The exterior orientation parameters were determined as follows. Firstly, bundle block adjustment was carried out using 20 ground control points and self-calibration of additional parameters using the program BLUH developed at the Institute of Photogrammetry and GeoInformation at Leibniz University Hannover. The effects of the systematic error correction were well below one pixel; cf. [Jacobsen et al., 2010] for details on the procedure and the results. In order to obtain the exterior orientation parameters in Table 4 that refer to a camera modelled to be free of distortion, a second bundle block adjustment was carried out, using the tie points from the first adjustment as ground control points and neglecting the additional parameters. Using the exterior orientation parameters described in this document should result in a back-projection error better than one pixel (RMS).

Strip	Image file	Projection Centres			Rotation Angles (ω: primary, x; φ: secondary, y; κ: tertiary, z)		
		X_0 [m]	Y_0 [m]	Z_0 [m]	ω [gon]	φ [gon]	κ [gon]
3	10030060.tif	496803.043	5420298.566	1163.983	2.50674	0.73802	199.32970
	10030061.tif	497049.238	5420301.525	1163.806	2.05968	0.67409	199.23470
	10030062.tif	497294.288	5420301.839	1163.759	1.97825	0.51201	198.84290
	10030063.tif	497539.821	5420299.469	1164.423	1.40457	0.38326	198.88310
4	10040081.tif	496558.488	5419884.008	1181.985	-0.87093	-0.36520	-199.20110
	10040082.tif	496804.479	5419882.183	1183.373	-0.26935	-0.63812	-198.97290
	10040083.tif	497048.699	5419882.847	1184.616	0.34834	-0.40178	-199.44720
	10040084.tif	497296.587	5419884.550	1185.010	0.81501	-0.53024	-199.35600
	10040085.tif	497540.779	5419886.806	1184.876	1.38534	-0.46333	-199.85010
5	10050103.tif	496573.389	5419477.807	1161.431	-0.48280	-0.03105	-0.23869
	10050104.tif	496817.972	5419476.832	1161.406	-0.65210	-0.06311	-0.17326
	10050105.tif	497064.985	5419476.630	1159.940	-0.74655	0.11683	-0.09710
	10050106.tif	497312.996	5419477.065	1158.888	-0.53451	-0.19025	-0.13489
	10050107.tif	497555.389	5419477.724	1158.655	-0.55312	-0.12844	-0.13636
25	10250130.tif	497622.784	5420189.950	1180.494	0.09448	3.41227	-101.14170
	10250131.tif	497630.734	5419944.364	1181.015	0.61065	2.54420	-97.84478
	10250132.tif	497633.024	5419698.973	1179.964	1.27053	1.62793	-97.23292
	10250133.tif	497628.317	5419452.807	1179.237	0.90688	0.83308	-98.72504
	10250134.tif	497620.954	5419207.621	1178.201	0.17675	1.27920	-101.86160
	10250135.tif	497617.307	5418960.618	1176.629	0.22019	1.47729	-101.55860

Table 4: Exterior orientation of the digital images of the Vaihingen Block.

2.1.3. DSM and True Orthophoto Mosaic

For participants interested in object detection we have provided a DSM and a true orthophoto generated from images. Both data sets are defined on the same grid with a ground resolution of 9°cm. The extents of the data are shown in Table°5. Note that a part of the area covered by the DSM grid does not contain any data. In the DSM, these void areas are marked by a height value of "-9999.0". In the true orthophoto, void areas are marked by grey levels of 0. In the mosaicking process, the radiometric resolution of the images was reduced to 8°bit. The DSM was generated with Trimble INPHO 5.3 software, using the modules MATCH-AT, MATCH-T DSM, SCOP++, and DTMaster. MATCH-T DSM applies a sequential multi-image matching procedure in several scales combining feature-based and least squares matching [Lemaire, 2008]. DSM results were moderately smoothed to reduce artifacts. Trimble INPHO OrthoVista was used to compute

a true orthophoto mosaic making use of adaptive feathering to get smooth transitions between adjacent images in the mosaic. In the DSM, small void areas were filled using a variant of nonlinear diffusion that is adaptive to height changes [Kosov et al., 2012].

	X_{min} [m]	Y_{min} [m]	X_{max} [m]	Y_{max} [m]	Height [pixels]	Width [pixels]	Number format
DSM	496274.985	5418933.165	498097.395	5420850.075	21300	20250	32 bit (float)
TOP	496274.985	5418933.165	498097.395	5420850.075	21300	20250	8 bit / band

Table 5: Area covered by the DSM and the true orthophoto mosaic (TOP) of the Vaihingen test site (centres of the corner pixels). The GSD of both data sets is 9°cm.

2.1.4. ALS Data

The Vaihingen test data set provided by DGPF also contains Airborne Laserscanner (ALS) data. The entire DGPF data set consists of 10 ALS strips acquired on 21 August 2008 by Leica Geosystems using a Leica ALS50 system with 45° field of view and a mean flying height above ground of 500°m. The average strip overlap is 30%, and the median point density is 6.7 points / m^2. Point density varies considerably over the whole block depending on the overlap, but in regions covered by only one strip the mean point density is 4 points / m^2. Multiple echoes and intensities were recorded. Due to the leave-on-conditions at the time of data acquisition, the number of points with multiple echoes is relatively low. The original point clouds were post-processed by strip adjustment to correct for systematic errors in georeferencing. In this process, object planes derived from the 8 cm DMC block were used as ground control, so that the georeferencing of the ALS data is consistent with the exterior orientation of the DMC images. As a result of the strip adjustment, the standard deviation derived from the median of absolute deviation in the overlap areas, σ_{MAD}, is σ_{MAD} = 2.9 cm. The test areas only overlap with five of the 10 strips, and only the overlapping areas of these four strips are provided. Each strip is provided in a separate file in las-format. Figure 5 gives an overview about the ALS data provided and the position of the test areas with respect to the ALS data. Table 6 shows which strips overlap with the individual test areas.
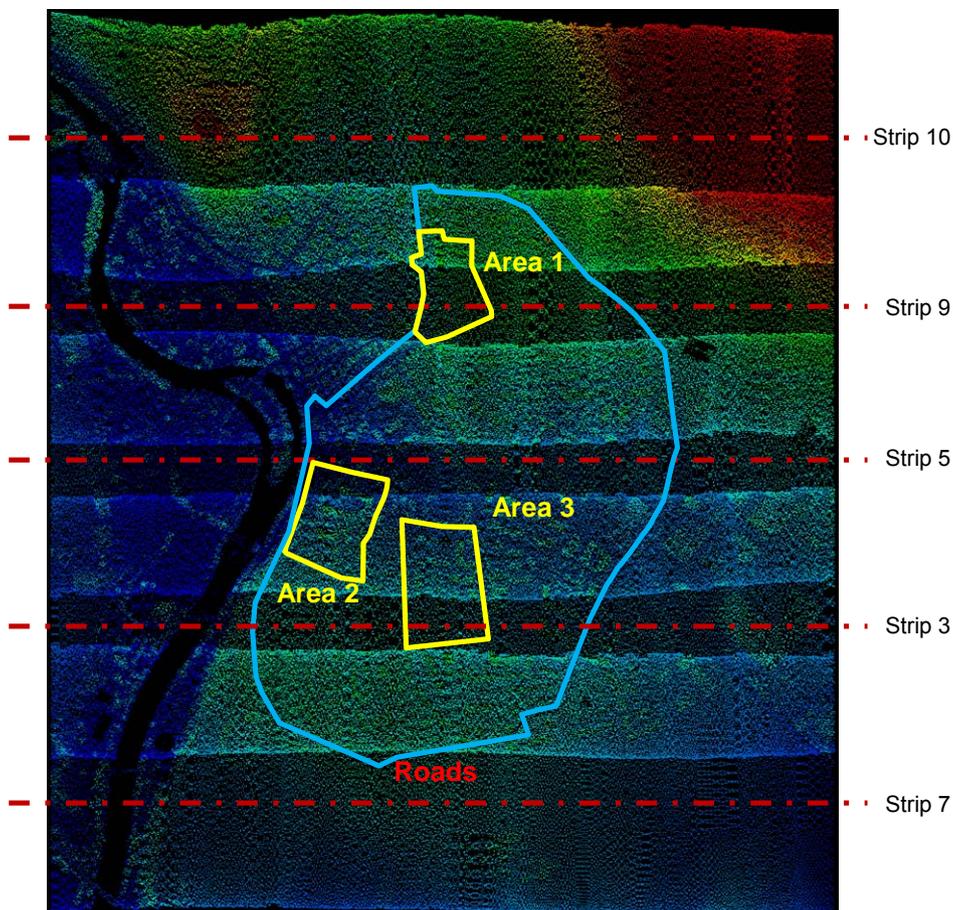
Figure 5: The ALS data of the Vaihingen test site.

In addition to the original ALS point cloud, a DSM is provided. This DSM was interpolated from the ALS point cloud with a grid width of 25 cm, using only the points corresponding to the last pulse. Figure 6 shows the DSM and the locations of the test areas within the DSM. Table 7 gives the extents of the DSM.

	Strip 3	Strip 5	Strip 7	Strip 9	Strip 10
Area 1	-	-	-	X	X
Area 2	X	X	-	-	-
Area 3	X	X	-	-	-
Roads	X	X	X	X	X

Table 6: Overview about the overlap between the test areas and the ALS data in the Vaihingen test data.
X: The strip test area overlaps with the strip; **-**: The test area does not overlap with the strip.

	X_{min}	Y_{min}	X_{max}	Y_{max}
DSM	496400.00	5418800.00	497850.00	5420400.00

Table7: Area covered by the ALS DSM of the Vaihingen test site (centres of the corner pixels).
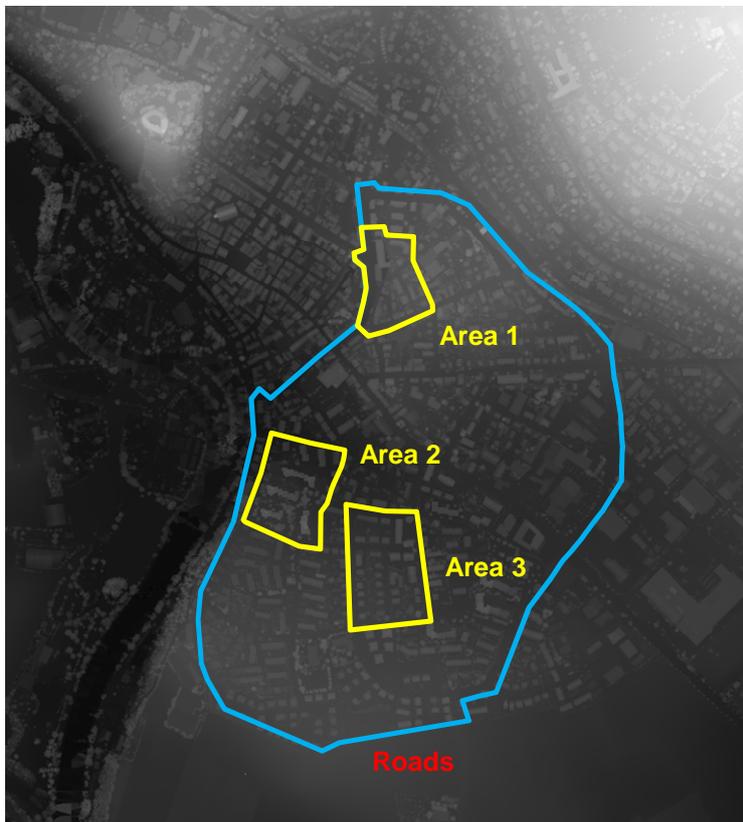
Figure 6: The ALS DSM data of the Vaihingen test site.

2.1.5. Conditions of Use

The Vaihingen test data are distributed subject to the following conditions:

1) **The data must not be used for other than research purposes.** Any other use is prohibited.

2) **The data must not be distributed to third parties.** Any person interested in the data may obtain them via ISPRS WG III/4.

3) Any scientific papers whose results are based on the Vaihingen test data must cite [Cramer, 2010] and must contain the following **acknowledgement**:
   "The Vaihingen data set was provided by the German Society for Photogrammetry, Remote Sensing and Geoinformation (DGPF) [Cramer, 2010]: http://www.ifp.uni-stuttgart.de/dgpf/DKEP-Allg.html."

4) The German Association of Photogrammetry, Remote Sensing and GeoInformation (DGPF) should be informed about any published papers whose results are based on the Vaihingen test by an e-mail to the DGPF Secretary. Currently, this is Eberhard Gülch (eberhard.guelch@hft-stuttgart.de).

Researchers using the Vaihingen test data for the evaluation of their object extraction techniques are encouraged to publish their results in the peer-reviewed journal of the DGPF, 'Photogrammetrie-Fernerkundung-Geoinformation' (PFG; http://www.dgpf.de/neu/pfg/general.htm). Please note that PFG also accepts papers written in English. PFG is indexed by the Science Citations Index Expanded.

2.2. Data Set 2: Downtown Toronto

2.2.1. Overview

This data set covers an area of about 1.45 km^2 in the central area of City of Toronto in Canada which was captured by the Microsoft Vexcel's UltraCam-D (UCD) camera and the Optech airborne laserscanner ALTM-ORION M. The "Downtown Toronto" data contains representative scene characteristics of a modern mega city in North America including a mixture of low- and high-story buildings with a wide variety of rooftop structures and street and road features. The "Downtown Toronto" data is divided into three scenes, 'Area 4', 'Area 5' and 'Entire Data'. The areas 'Area 4' and 'Area 5' should be used for conducting comparative performance test of object extraction and building reconstruction algorithms, while the 'Entire Data' should be used for testing road detection algorithm (see Figures 7 and 8). The coordinates of the four corners of the entire test area are shown in Table 8.

- **Area 4:** This area contains a mixture of low and high story buildings, showing various degrees of shape complexity containing rooftop structures and rooftop furniture. The scene also contains different urban objects including cars, trees, street furniture, roads and parking lots (Figure 8a).

- **Area 5:** This area represents a typical example of a cluster of high-rise buildings in a modern mega city in North America. The scene contains shadows casted by high-rise buildings, under which diverse urban objects (e.g., cars, street furniture, and roads) can be found (Figure 8b).

Figure 7: The test area in downtown Toronto

2.2.2. Digital Aerial Images

Digital Aerial Images taken by UltraCam-D cover the downtown of Toronto which was operated and processed by FBS (First Base Solutions) company located in the Greater Toronto Area in Canada (http://www.firstbasesolutions.com/). The data consist of three overlapping strips with 30% side lap and 60% forward overlap. The total number of the images is 13 and the exterior orientation parameters are provided.

The image size is 7500 × 11500 pixel and the pixel size is 9 µm. Table 9 gives a summary of the configuration of the block, whereas Table 10 gives an overview of the digital aerial images. The image configuration is depicted in Figure 9.
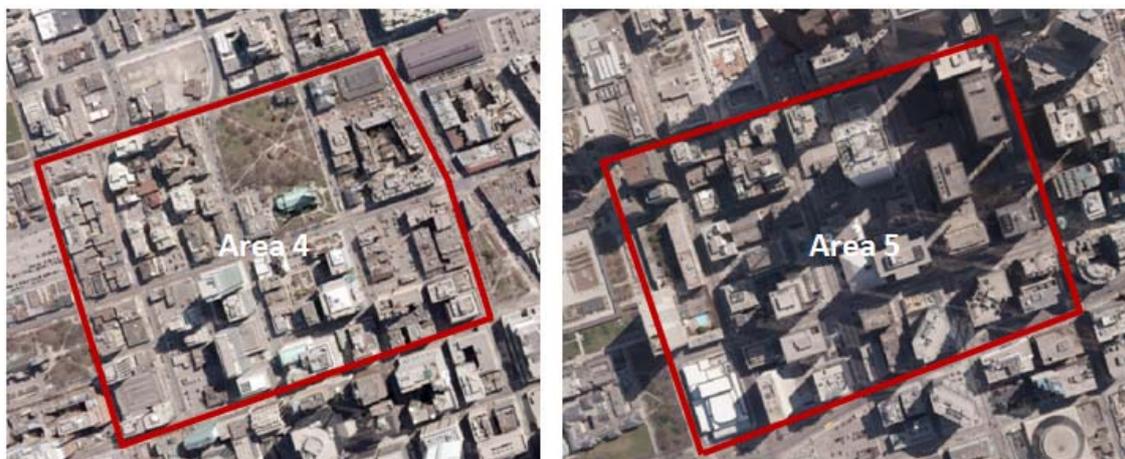
Figure 8: The two test sites in Toronto. Left: Area 4, Right: Area 5
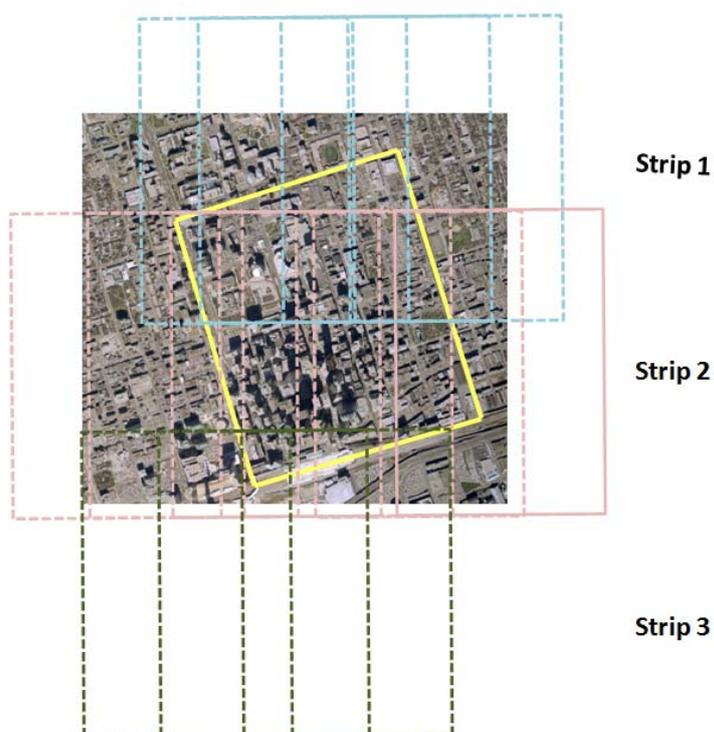
Figure 9: Image configuration for the downtown Toronto test site.

Upper Left		Upper Right		Lower Left		Lower Right	
E(m)	N(m)	E(m)	N(m)	E(m)	N(m)	E(m)	N(m)
629955	4834885	631010	4835253	630355	4833650	631380	4833954

Table 8: Coordinates of the corners of the entire test area of the Toronto test site (cf. Figure 7).

Camera	Focal length	Flying height above Ground	Forward overlap	Side lap	GSD	Spectral bands	Radiometric resolution
UltraCam D	101.4 mm	1600 m	60 %	30 %	15 cm	R-G-B	8 bit

Table 9: Flight parameters of the Toronto block.

	Strip 1	Strip 2	Strip 3
Area 4	03557*	03753, 03755, 03757*	-
Area 5	-	03749*, 03751, 03753	-
Downtown Toronto	03553, 03555, 03557, 03559*	03747*, 03749*, 03751, 03753, 03755, 03757*	03945*, 03947*, 03949*

Table 10: Overview about the images of the Toronto Block. The asterisk (*) means that the area is only partially visible in that image.

2.2.3. Orientation Parameters

Table 11 shows the specification of the interior orientation of the camera used for the Toronto test area, whereas Table 12 gives the exterior orientation parameters of each image. All the coordinate systems and the sensor parameters describing the "Downtown Toronto" images follows the definitions used in Section 2.1.3 ("Vaihingen" data). The planimetric coordinates of the object coordinate system refer to WGS84 and UTM Zone 17 North; the heights are geodetic heights.

Camera	Strip	file coordinate system			camera coordinate system			pixel size Δ [mm]
		row_{PP} [pixel]	col_{PP} [pixel]	f [pixel]	x_{PP} [mm]	y_{PP} [mm]	f [mm]	
UltraCam D	1,3	3730	5750	11266.67	0.000	0.180	101.400	0.009
	2	3770	5750	11266.67	0.000	-0.180	101.400	0.009

Table 11: Interior orientation of the digital images of the Toronto test site.

Strip	Image file	Projection Centres			Rotation Angles (ω: primary, x; φ: secondary, y; κ: tertiary, z)		
		X_0 [m]	Y_0 [m]	Z_0 [m]	ω [gon]	φ [gon]	κ [gon]
1	03553.tif	630203.843	4835116.640	1635.515	-0.09322	-0.10489	100.40389
	03555.tif	630593.522	4835117.251	1634.227	-0.07822	-0.15100	100.31978
	03557.tif	630982.549	4835118.458	1633.674	-0.09967	-0.15833	100.05244
	03559.tif	631372.022	4835118.729	1635.037	-0.08589	-0.20500	100.01922
2	03747.tif	629622.532	4834069.618	1635.711	0.03367	0.25544	-99.54756
	03749.tif	630011.621	4834067.266	1633.155	0.04494	0.20707	-100.09660
	03751.tif	630401.052	4834064.940	1631.818	0.04053	0.17273	-100.12400
	03753.tif	630789.997	4834064.034	1632.821	0.03059	0.12878	-100.16312
	03755.tif	631179.232	4834062.777	1636.415	0.02118	0.08621	-100.34622
	03757.tif	631567.432	4834062.247	1640.148	0.04880	0.02709	-100.21350
3	03945.tif	630003.813	4833037.773	1642.041	-0.10900	-0.05267	100.69900
	03947.tif	630393.759	4833038.863	1639.798	-0.12856	-0.10867	100.35722
	03949.tif	630781.715	4833039.040	1636.504	-0.10167	-0.15244	99.89067

Table 12: Exterior orientation of the digital images of the Toronto Block.

2.2.4. ALS Data

In addition to the UltraCam-D images, the "Downtown Toronto" datasets also provides ALS data acquired by Optech (http://www.optech.ca/). Optech flew over the "Downtown Toronto" area and acquired ALS data using Optech's ALTM-ORION M in February 2009 with the aircraft speed of 120 knots at the flying altitude of 650 m. The ALTM ORION M operates at a wavelength of 1064 nm (Near Infrared) and scans the underlying topography with a scan width of 20 degrees and the scan frequency of 50 Hz. The reflected echoes were digitized at a sampling rate of 100 kHz. The data set consists of 6 strips and point density is approximately 6.0 points/m^2. Figure 10 shows the ALS data over the "Downtown Toronto" region. The ALS data provided is formatted in ASPRS (American Society of Photogrammetry and Remote Sensing)'s LAS 1.3 format and refers to the same coordinate system as the orientation parameters of the UltraCam-D images.

In addition to the original ALS point cloud, a digital surface model (DSM) is provided. This DSM was interpolated from the ALS point cloud with a grid width of 25 cm, using only the points corresponding to the last pulse. Figure 11 shows the DSM. Table 13 gives the extents of the DSM.

	X_{min}	Y_{min}	X_{max}	Y_{max}
DSM	629939.245	4833596.245	631413.495	4835248.995

Table 13: Area covered by the ALS DSM of the Toronto test site.
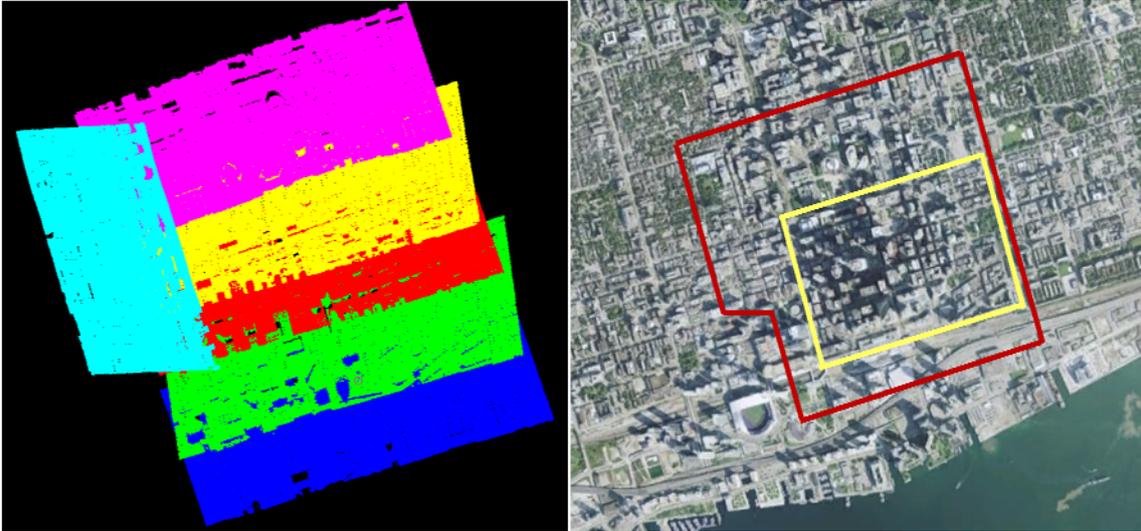
Figure 10: ALS data for the Toronto test site. Left: ALS strips. Right: coverage of ALS strips (red) and of test site (yellow)
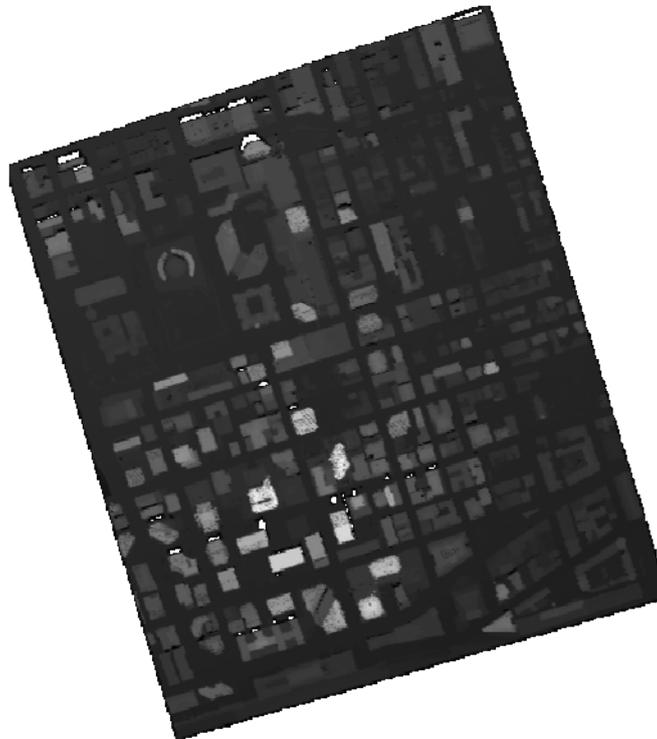
Figure 11: The DSM data of the Toronto test site.

2.2.5. Conditions of Use

The Toronto test data are distributed subject to the following conditions:

1) The data **must not be used for other than research purposes.** Any other use is prohibited.

2) The data **must not be used outside the context of this test project**, in particular while the project is still on-going (i.e. until September 2012). Whether the data will be available for other research purposes after the end of this project is still under discussion.

3) **The data must not be distributed to third parties.** Any person interested in the data may obtain them via ISPRS WG III/4.

4) The data users should include the following acknowledgement in any publication resulting from the datasets:

   "*The authors would like to acknowledge the provision of the Downtown Toronto data set by Optech Inc., First Base Solutions Inc., GeoICT Lab at York University, and ISPRS WG III/4.*"

2.3. Location of the Data Files

The data is hosted on an ftp-server at the University of Twente, Faculty ITC (The Netherlands). Detailed instructions on how to access the data will be sent to the individual participants by email. In order to enable a proper registration of participants, a questionnaire form needs to be filled in. That form is accessible via http://www.itc.nl/ISPRS_WGIII4/tests_datasets.html . Especially the terms of use need to be acknowledged before we are allowed to provide participants with the data.

2.3.1. Data Set 1: Vaihingen

The Vaihingen test data set is located in directory `Vaihingen` found in the root directory of the ftp server. It has a total download size of approx. 17 GB. The data can be found in the following sub-directories:

- `Vaihingen/images` contains the image data files (`name.tif`, where `name` is the image identifier used in this text, e.g. `10040082.tif` for image 10040082) and the files containing the orientation parameters (`daporo.dat, daporp.dat`)

- `Vaihingen/ALS` contains the ALS data in las-format. There is one las-file per strip named `Vaihingen_Strip_NN.LAS`, where `NN` is the two-digit strip number.

- `Vaihingen/DSM` contains the Digital Surface Models as Geo-TIFF-files with 32 bits per (`float`) height values. The DSM from ALS is named `DSM_25cm_ALS.tif`, whereas the DSM from matching is called `DSM_09cm_matching.tif`. In addition, there are two World Files, `DSM_25cm_ALS.tfw` and `DSM_09cm_matching.tfw`, respectively, containing the georeferencing.

- `Vaihingen/Ortho` contains the true orthophoto mosaic as an 8 bit RGB GeoTIFF file named `TOP_Mosaic_09cm.tif`. The World File is `TOP_Mosaic_09cm.tfw`.

2.3.2. Data Set 2: Downtown Toronto

The Downtown Toronto test data set is located in directory `Toronto` found in the root directory of the ftp server. It has a total download size of approx. 5.5 GB. The data can be found in the following sub-directories:

- `Toronto/images` contains the image data files (`name.tif`, where `name` is the image identifier used in this text, e.g. `03947.tif` for image 03947) and the files containing the orientation parameters (`daporo.dat, daporp.dat`)

- `Toronto/ALS` contains the ALS data in las-format. There is one las-file per strip named `Toronto_Strip_NN.LAS`, where `NN` is the two-digit strip number.

- `Toronto/DSM` contains the Digital Surface Model as Geo-TIFF-files with 32 bits per (`float`) height value. There is on Geo-TIFF file named `Toronto_DSM_25cm_ALS.tif` and in addition a World File `Toronto_DSM_25cm_ALS.tfw` containing the georeferencing.

3. Tasks

3.1. Urban Object Extraction

The participants should carry out object detection in the test areas. Participants should deliver detection results for one or more of the following object classes:

- **Buildings:** The results can be delivered as closed 2D or 3D polylines describing the building outlines in DXF format, in the form of a binary building mask as a geocoded Tiff file, or in the form of a building label image as a geocoded Tiff file.

- **Roads:** The results can be delivered in one of two ways:

    1. As 2D or 3D polylines describing the road centre lines in DXF format, optionally with a width parameter. Polygons must be split at crossroads, i.e. an intersection point of two or more road axes at a crossroads must be one of the end points of all polylines emanating from the crossroads.

    2. As 2D or 3D polylines describing the road edges in DXF format. Again, the polylines must be split at crossroads. There should be one DXF layer for each road segment, and the two polylines corresponding to the two edges of a road segment must be assigned to the same layer (Figure 13).

- **Trees:** The results can be delivered as closed 2D or 3D polylines describing the outlines of tree crowns in DXF format, in the form of a binary tree mask as a geocoded Tiff file, or in the form of a tree label image as a geocoded Tiff file.

- **'Artificial' ground other than road:** This class contains all areas on the ground that do not correspond to roads but are covered by materials such as asphalt that are typically used for paving roads. In particular, it contains parking lots, pavements, inner courtyards and driveways (if paved). The results can be delivered as closed 2D or 3D polylines describing the outlines of such areas in DXF format, in the form of a binary mask as a geocoded Tiff file, or in the form of a label image of such areas as a geocoded Tiff file.

- **'Natural' ground covered by vegetation**: This class contains any areas on the ground covered by vegetation other than trees. In particular, it contains lawn and low bushes. The results can be delivered as closed 2D or 3D polylines describing the outlines of such areas in DXF format, in the form of a binary mask as a geocoded Tiff file, or in the form of a label image of such areas as a geocoded Tiff file.

- **Cars**: Any moving or static cars inside the test areas should be extracted by participants interested in detecting cars. The results can be delivered either as closed (rectangular) 2D or 3D polylines in DXF format, each describing the outlines of a car, or in the form of a label image of cars as a geocoded Tiff file. For each of the test areas, one image is defined from which the reference is generated (Table 14).

a) Road modelled by their centre lines   b) Roads modelled by their edges
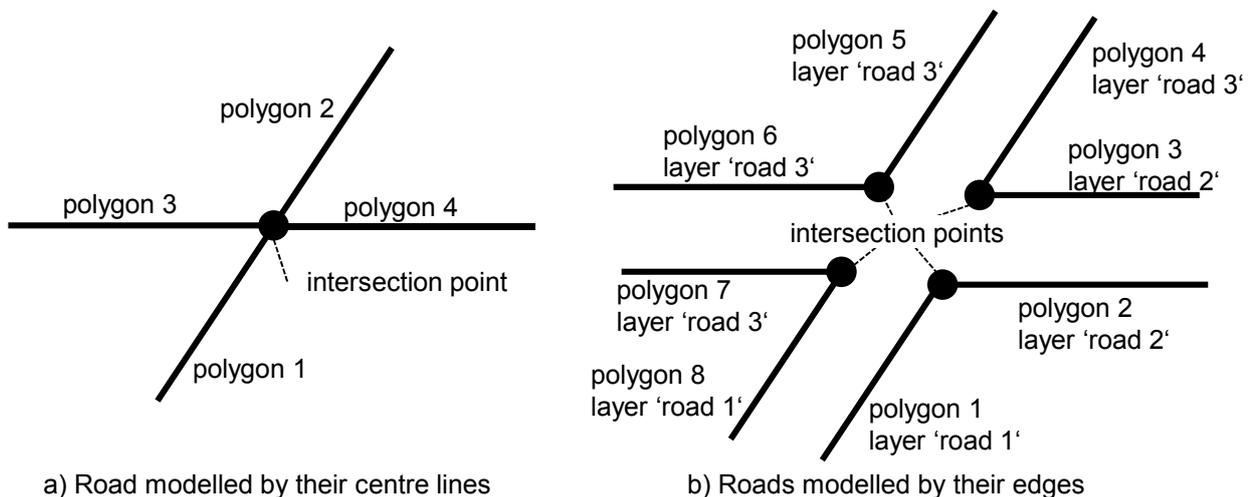
Figure 13: Two alternatives for delivering the results of road extraction.

For the Vaihingen data set, road extraction should be carried out for the test area 'Roads', whereas the other objects should only be detected in the three test areas 'Area 1', Area 2', and 'Area 3'. For the Downtown Toronto data set, road data extraction should be performed for the entire test area, whereas trees and nature should be extracted in the test area 'Area 4'. The other objects should be extracted in the two test areas 'Area 4' and 'Area 5'.

All the results must be delivered in the object coordinate system given by the respective test area. Only the planimetry will be taken into account for the evaluation. Participants may choose to extract a single object class or any subset of the classes defined above. Object extraction techniques capable of delivering multiple object classes simultaneously are encouraged, as they correspond well with the topic of ISPRS WG III/4,

"Complex Scene Analysis". In any case, **results should be submitted in separate files for each object class** in order to facilitate the evaluation of these results. Along with the results of object detection, a report on how these results were achieved along with references to detailed descriptions of the methodology.

Vaihingen	Area	Area 1	Area 2	Area 3	Roads
	Image	10040083	10050105	10050104	-
Downtown Toronto	Area	Area 4		Area 5	
	Image	03755		03751	

Table 14: Images to be used for car detection.

The results submitted by the test participants will be evaluated by the test organizers based on reference data. The reference data for Vaihingen were generated by photogrammetric plotting. The basis is the digital map generated by RAG [Spreckels et al., 2010]. It was augmented by additional object classes by Ms D. Müller B.Sc. at the Institute of Photogrammetry and GeoInformation at Leibniz University Hannover, Germany. The reference data for Downtown Toronto were provided by City of Toronto, First Base Solutions and York University's GeoICT Lab. City of Toronto provides vectors of building footprints, road central lines and boundaries; First Base Solutions provides its own product of 3D building rooftop models produced for the purpose of GeoBrowsing. York University is responsible for quality control of all kinds of reference data and evaluation of participants' object detection.

All objects except roads will be evaluated by a comparison of label images based on the technique described in [Rutzinger et al., 2009] that provides **completeness**, **correctness**, and **quality** of the results both on a per-object and on a per-area level. For participants delivering polygons, the 2D RMS error of the object outlines of the correct objects will be determined as well. For roads, the evaluation technique described in [Wiedemann & Ebner, 2000] based on a buffer method will be used.

3.2. 3D Building Reconstruction

Participants shall generate detailed 3D models of the building roofs in the test areas. The goal of this task is to derive a complete, correct, and accurate segmentation of the roof planes in the provided data. The level of detail should correspond to LoD2 of the CityGML standard [Gröger et al., 2008]. That is, the roof models should contain all the major roof structures, including even small dormers, but no roof overhangs, no façade details, and no details such as balconies are to be modeled. The results shall be submitted as DXF files containing closed 3D polygons corresponding to the boundaries of the reconstructed roof planes in the object coordinate system given by the respective test area. If complete building models, including walls and floors, are delivered, roof polygons should be marked by assigning them to a separate layer 'roof'. The reference for Vaihingen was generated by photogrammetric plotting carried out by the SIRADEL company in France (www.siradel.com), following the guidelines used by RAG in Area 1 [Spreckels et al., 2010]. The reference for Downtown Toronto (LoD2) was generated by photogrammetric plotting by York University's GeoICT Lab (www.geoict.yorku.ca) based on a LoD1 model generated by the City of Toronto. The evaluation will consist of an analysis of the quality of the segmentation and an analysis of the geometrical errors of the submitted models. The analysis of the quality of the segmentation will be based on a comparison of roof plane label images and will focus on missed or oversegmented planes, and on topological errors, i.e. missing or incorrect neighbourhood relations as a consequence of segmentation errors. The geometrical error will be evaluated by determining the RMS errors of building roof vertices (only for roof planes correctly segmented) and of an overall analysis of the height differences between the submitted models and the reference.

3.3. Submission of Results

Submission of results is again done via ftp. Detailed instructions can be found on
http://www.itc.nl/ISPRS_WGIII4/tests_datasets.html . We ask you to stick to the file formats as indicated in Sections 3.1 and 3.2. For any technical questions concerning the submission please contact the person as indicated on the website.

3.4. Publications

Due to its success, the test is still ongoing and will remain so in the future. Participants can submit their results for evaluation at any time; they will receive the results of the evaluation within one or two days.

Evaluation results are also made available online:
http://www.itc.nl/ISPRS_WGIII4/ISPRSIII_4_Test_results/tests_datasets_results_main.html.

The evaluation results may be used by the participants for any scientific publication. A **Theme Issue on Urban Object Detection and 3D Building Reconstruction of the ISPRS Journal of Photogrammetry and Remote Sensing** is in the reviewing phase; first papers are already available online via the journal's web site. The planned date of publication of the theme issue is summer 2014.

4. Timetable

7 March 2011	Announcement of the benchmark, data distribution by ISPRS WG III/4
~~30 September 2011~~	~~Deadline for submitting results by the participants who want to submit full papers for the peer-reviewed track of ISPRS Commission III at the ISPRS Congress in Melbourne~~
~~30 October 2011~~	~~Participants having submitted their results by 30 September are informed about the evaluation of their results~~
~~28 November 2011~~	~~Deadline for submitting full papers for the peer-reviewed track of ISPRS Commission III at the ISPRS Congress in Melbourne~~
~~31 May 2012~~	~~Final deadline for submitting results by the participants~~
~~30 June 2012~~	~~Participants are informed about the evaluation of their results~~
~~24 August–3 September 2012~~	~~ISPRS Congress in Melbourne~~
~~15 May 2013~~	~~Final deadline for submitting results by the participants for inclusion in overview paper~~
~~31 May 2013~~	~~Deadline for paper submission to the theme issue of the ISPRS Journal~~
Summer 2014	Planned publication of the theme issue of the ISPRS Journal
Ongoing	Evaluation of results submitted by the participants

References

Cramer, M., 2010. The DGPF test on digital aerial camera evaluation – overview and test design. Photogrammetrie – Fernerkundung – Geoinformation 2(2010):73-82.

Gröger, G., Kolbe, T. H., Czerwinski, A., Nagel, C. 2008. OpenGIS city geography markup language (CityGML) encoding standard, Version 1.0.0, OGC Doc. No. 08-007r1, Open Geospatial Consortium; URL (accessed 23 December 2010): http://www.opengeospatial.org/standards/citygml

Haala, N., Hastedt, H., Wolf, K., Ressl, C., Baltrusch, S., 2010. Digital photogrammetric camera evaluation – generation of digital elevation models. Photogrammetrie – Fernerkundung – Geoinformation 2(2010):99-115.

Jacobsen, K., Cramer, M., Ladstädter, R., Ressl, C., Spreckels, V., 2010. DGPF project: evaluation of digital photogrammetric camera systems – geometric performance. Photogrammetrie – Fernerkundung – Geoinformation 2(2010):83-97.

Kosov, S., Rottensteine, F., Heipke. C., Leitloff, J., S. Hinz, S., 2012. 3D classification of crossroads from multiple aerial images using Markov Random Fields. Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 39(B3), pp. 479-484.

Lemaire, C., 2008. Aspects of the DSM production with high resolution images. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 37(B4), pp. 1143-1146.

Rutzinger, M., Rottensteiner, F. Pfeifer, N., 2009. A comparison of evaluation techniques for building extraction from airborne laser scanning. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 2(1):11-20.

Spreckels, V., Syrek, L., Schlienkamp, A., 2010. DGPF project: evaluation of digital photogrammetric camera systems - stereoplotting. Photogrammetrie – Fernerkundung – Geoinformation 2(2010):117-130.

Wiedemann, C., Ebner, H., 2000. Automatic completion and evaluation of road networks. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 33 (B3/2), pp. 979-986.

Contact

For further information, contact any of the officials of ISPRS WG III/4

Franz Rottensteiner: rottensteiner@ipi.uni-hannover.de

Gunho Sohn: gsohn@yorku.ca

Markus Gerke: gerke@itc.nl

Jan Dirk Wegner: jan.wegner@geod.baug.ethz.ch

You can also visit the WG web site: http://www.commission3.isprs.org/wg4/

07 January 2013