

## DETECTION OF TREE CROWNS BASED ON RECLASSIFICATION USING AERIAL IMAGES AND LIDAR DATA

S. Talebi <sup>a,\*</sup>, A. Zarea <sup>b</sup>, S. Sadeghian <sup>c</sup>, H. Arefi <sup>d</sup>

<sup>a</sup> Geomatics Engineering Faculty, Tafresh State University, Tafresh, Iran

<sup>b</sup> Geodesy and Geomatics Engineering Faculty, K.N. Toosi University of Technology, Tehran, Iran

<sup>c</sup> Geomatics college of National Cartographic Center (NCC), Tehran, Iran, P.O.Box: 13185-1684

<sup>d</sup> Remote Sensing Technology Institute, German Aerospace Center (DLR), 82234, Wessling, Germany

### Commission III, WG III/4

**KEY WORDS:** nDSM, Shadow Index, NDVI, Unsupervised, Hierarchical, LiDAR, Reclassification, Tree Detection

#### ABSTRACT:

Tree detection using aerial sensors in early decades was focused by many researchers in different fields including Remote Sensing and Photogrammetry. This paper is intended to detect trees in complex city areas using aerial imagery and laser scanning data. Our methodology is a hierarchical unsupervised method consists of some primitive operations. This method could be divided into three sections, in which, first section uses aerial imagery and both second and third sections use laser scanners data. In the first section a vegetation cover mask is created in both sunny and shadowed areas. In the second section Rate of Slope Change (RSC) is used to eliminate grasses. In the third section a Digital Terrain Model (DTM) is obtained from LiDAR data. By using DTM and Digital Surface Model (DSM) we would get to Normalized Digital Surface Model (nDSM). Then objects which are lower than a specific height are eliminated. Now there are three result layers from three sections. At the end multiplication operation is used to get final result layer. This layer will be smoothed by morphological operations. The result layer is sent to WG III/4 to evaluate. The evaluation result shows that our method has a good rank in comparing to other participants' methods in ISPRS WG III/4, when assessed in terms of 5 indices including area base completeness, area base correctness, object base completeness, object base correctness and boundary RMS. With regarding of being unsupervised and automatic, this method is improvable and could be integrate with other methods to get best results.

### 1. INTRODUCTION

Detection and classification of objects on earth were and still are important fields for researchers in different majors including Remote Sensing and Photogrammetry (Rottensteiner et al., 2011). As emerging new sensors like laser scanners, developing Photogrammetry field, and utilizing digital cameras, methods of detection and classification are got into new era. High resolution and high spectral digital cameras have lead researchers to develop and introduce new indices to detect a variety of objects on earth. LiDAR data give 3D coordinates of points directly, that this ability makes it a simple task to differentiate between smooth and rough planes. Smooth planes usually designate man-made objects and rough planes mostly designate natural grounds. Because LiDAR is an active sensor it has no problem dealing with shadow areas, while shadow areas are challenging in aerial images. At high resolution aerial images boundary of objects like Buildings is clearly notable, but in LiDAR data there are some problems in detecting such boundaries. Considering advantages and disadvantages both LiDAR and aerial imagery it seems integrating these two data sources is the best option (Rottensteiner et al., 2008). Tree detection in complex city scenes because of existing high buildings is a more difficult task than tree detection in plains and cities with low buildings. There are lots of methods and algorithms in detection and classification field but it is not possible to compare those together. This is because of lack of bench mark data sets. In other hand most of algorithms and methods have tested in different data sets. To overcome this problem and making it easier to compare methods together a

working group established in ISPRS, named WG III/4. This WG grants a bench mark data set to participants and encourages them to test their methods on this data and send the results to WG for evaluation (F. Rottensteiner et al., 2011). In continue some of related works about tree detection are mentioned. We separated WG III/4 participants' related works from the others.

#### 1.1 Related works

Tucker introduced Normalized Difference Vegetation Index (NDVI) on the basis of plants reflection properties in Red (R) and Infrared (IR) bands. This index is used to detect vegetation cover (Tucker, 1979). Ono introduced a new Shadow Index (SI) in 2007 (Ono, 2007), that Grigillio in his article used this index to detect buildings (Grigillio et al., 2011). Cheuk-Yan in his article enhanced bands which were in shadow areas using Gamma Correction (GC) and Linear Correlation Correction (LCC). According on his article detection of trees after enhancement was much easier, because enhanced areas were spectrally more near to non-shadow areas (Cheuk-Yan Wan et al., 2012). Cai introduced Shadow Index based on Hue Saturation Intensity (HSI) colour space and used difference, ratio and normalized difference to detect shadow areas, simultaneously he used NDVI to improve the results (Dong Cai et al., 2010). Sarabandi introduced a new transformation for shadow detection based on Color Invariant Indices (CII) (Sarabandi et al., 2006).

\* Email: [eng.siamaktalebi@gmail.com](mailto:eng.siamaktalebi@gmail.com)

Mobile Phone: +98 - 9149718713

## 1.2 WG III/4 participants' works

**1.2.1 A. Moussa, Canada (CAL):** This method has used the combination of ALS point cloud and orthophoto data. In this method objects are classified into building, tree and ground segments using a rule-based segmentation. Employing spectral data obtained from aerial images the segmentation result is improved. At last morphological operations are used to smooth labeled image (Moussa & El-Sheimy, 2012).

**1.2.2 J. Niemeyer, Germany (HAN):** This method uses a supervised ALS point's classification on the basis of Conditional Random Fields (CRF), which employs a statistical model of context (Niemeyer et al., 2011). At the end the result label image is smoothed by morphological operations.

**1.2.3 D. Grigillo and U. Kanjir, Slovenia (LJU):** ALS points are used to calculate DTM, DSM and nDSM. Then a mask is used to obtain off-train objects. Buildings and trees are separated by NDVI. At last buildings boundaries are obtained by Hough Transform (Grigillo et al. 2011).

**1.2.4 Yao, Germany (TUM):** In this method both aerial images and range data are used in supervised classification. To obtain training data 10% of area is digitized manually.

**1.2.5 Q. Zhan, China (WHU):** Images are ortho-rectified using DTM. ALS points, spectral and range data are used in a supervised classification.

## 2. INPUT DATA

Our input data in this article are IR, R and G bands from aerial images and ALS point clouds. IR, R and G bands have wave lengths as below:

- IR covering from 0.79 to 89  $\mu\text{m}$ ,
- R covering from 0.61 to 0.68  $\mu\text{m}$  and
- G covering from 0.50 to 0.59  $\mu\text{m}$ .

ALS point clouds have 4 kinds of data including: First Pulse Range (FPR), Last Pulse Range (LPR), First Pulse Intensity (FPI) and Last Pulse Intensity (LPI). In this article we would use FPR.

## 3. METHODOLOGY

Fig. 1 shows a flowchart for our methodology. This method can be divided into 3 sections. First section uses aerial images and both two other sections use ALS point clouds as input data. In the first section NDVI and SI indices are calculated using aerial image bands. Using a ratio between these indices we would obtain a new index, called ESI. This index shows vegetation in shadow areas with more values per pixel. This index and NDVI are reclassified into some classes using equal intervals. Then a linear equation is wrote between these indices, which this equation causes vegetation cover whether in sunny areas or in shadow areas get closer to top classes. At the end of this section a rule-based function is used to obtain the first section result layer. The problem with this section is the fact that the result layer includes grasses and bushes which are not desirable. In two other sections we would try to overcome these problems. In the second section a RSC is calculated using DSM. This index could eliminate entities which have a low value e.g. grasses. In the third section employing DSM and morphological operations we would obtain DTM, going ahead nDSM would be calculated using DSM and DTM. At the end of this section a rule-based

function is used to reach the last result layer. The third section has ability to separate vegetation with low height, e.g. bushes, from higher vegetation, e.g. trees.

At the end we have three result layers from three sections. Using a point-wise multiplication of three last layers the total last result layer would be obtained. Morphological operations could be used to smooth the result layer. In the following we would explain each section with more details.

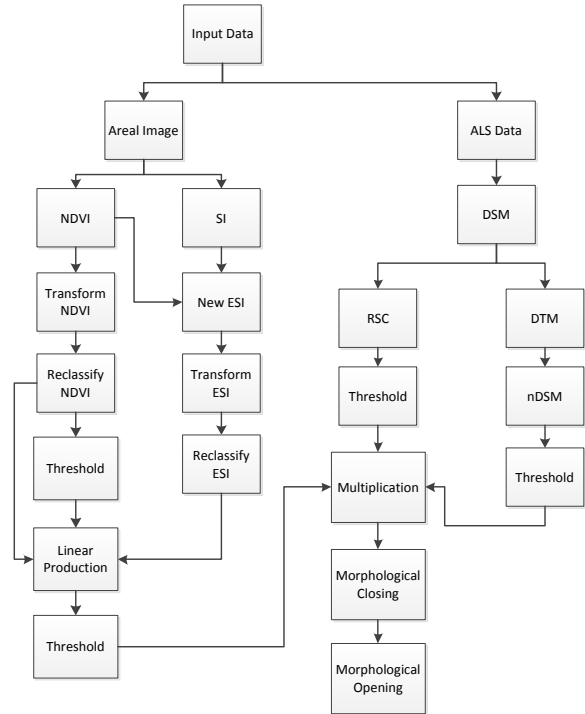


Figure . Representation of methodological workflow in brief

### 3.1 First section

**3.1.1 NDVI:** Using IR and R bands NDVI is calculated:

$$NDVI = \frac{IR - R}{IR + R} \quad (1)$$

Where

$IR$  = infrared band

$R$  = red band

$NDVI$  = normalized difference vegetation index

In this index brighter pixels show vegetation cover.

**3.1.2 The New Enhanced Shadow Index:** Ono, 2007 introduced SI using IR, R and G bands:

$$Normalized_R = \frac{R}{R + G + B} \quad (2)$$

$$SI = Normalized_R - R$$

Where

$B$  = blue band

$R$  = red band

$G$  = green band  
 $SI$  = shadow index

In this article we use this index by briefly changing:

$$Normalized_{IR} = \frac{IR}{IR + R + G} \quad (3)$$

$$SI = IR - Normalized_{IR}$$

Now by dividing NDVI on SI a new index is introduced. We call this index Enhanced Shadow Index (ESI). In this index higher values are showing vegetation cover in shadow areas:

$$ESI = \frac{NDVI}{SI} \quad (4)$$

Where  $ESI$  = Enhanced Shadow Index

**3.1.3 Reclassify:** ESI and NDVI would be reclassified into C classes with equal intervals, after scaling to a consistent scale:

$$\begin{aligned} ESI &= Transform(ESI) \\ NDVI &= Transform(NDVI) \\ Re\ ESI &= Re\ classify(ESI) \\ Re\ NDVI &= Re\ classify(NDVI) \end{aligned} \quad (5)$$

Where  $Re\ ESI$  = reclassified ESI  
 $Re\ NDVI$  = reclassified NDVI

**3.1.4 Linear Production:** Now employing a linear production we would get a new layer. In this layer vegetation cover existing in sunny and shadow areas is in high classes:

$$\begin{aligned} D &= Re\ NDVI > X \\ LP &= A \times Re\ NDVI + B \times Re\ ESI + D \end{aligned} \quad (6)$$

Where  $A$ ,  $B$  and  $X$  = consistent values which are obtained by practice  
 $D$  = represents classes of ReNDVI which are ensured as trees  
 $LP$  = linear production

To make clear Eq. 6 we represent Fig.2 with a numeric example:

ReNDVI			ReESI			ReNDVI>18			LP		
14	16	20	17	20	16	0	0	20	16	19	37
25	17	21	15	12	13	25	0	21	43	14	36
17	15	13	23	18	19	0	0	0	21	17	17
0.3*ReNDVI			0.7*ReESI			D			LP		

Figure . This figure shows the way of combining ReNDVI and ReESI

**3.1.5 First Section's Result Layer:** At the end of this section a threshold used to get the last result:

$$IMRES = LP > Y \quad (7)$$

Where  $Y$  = a consistent value which is obtained practically

### 3.2 Second Section

**3.2.1 Slope:** In this section using DSM, which is obtained using ALS point cloud, slope layer is calculated (Ritter, P. 1987):

$$\begin{aligned} Slope_{we} &= \frac{Z_3 - Z_1}{2 \times d} \\ Slope_{sn} &= \frac{Z_2 - Z_4}{2 \times d} \end{aligned} \quad (8)$$

Where  $Z_1, Z_2, Z_3$  and  $Z_4$  = elevations in 4-conectivity neighbourhood based on Fig. 3  
 $d$  = grid interval  
 $Slope_{we}$  = west-east slope  
 $Slope_{sn}$  = south-north slope

$Z_5$	$Z_2$	$Z_6$
$Z_1$	$Z_0$	$Z_3$
$Z_8$	$Z_4$	$Z_7$

Figure . Representation of number of pixels in neighbourhood

For computing total slope Eq. 9 is used (Ritter, P. 1987):

$$Slope = \sqrt{Slope_{sn}^2 + Slope_{we}^2} \quad (9)$$

**3.2.2 Rate of Slope Change:** Zhilin Li, 2005 in his book wrote Eq. 17 to calculate RSC (Zhilin Li et al., 2005), but here we computed this layer in other way:

$$RSC_0 = MAX \left( \frac{Slope_j - Slope_0}{\sqrt{2d}} \right), \text{ for } j=5, 6, 7, 8 \quad (10)$$

Where  $d$  = grid interval  
 $RSC_0$  = rate of slope change in pixel number 0

In this article for this purpose firstly slope layer is computed using DSM, and then these equations are repeated on slope layer to get RSC:

$$\begin{aligned} Slope &= Slope(DSM) \\ RSC &= Slope(Slope) \end{aligned} \quad (11)$$

**3.2.3 Section Two Result Layer:** After computing RSC a threshold is used to eliminate areas with low RSC:

$$SLOPRES = RSC > Z \quad (12)$$

Where  $Z$  = threshold value

### 3.3 Third Section

**3.3.1 nDSM:** In this section our aim is to separate entities with a specific height from other entities. For this purpose DTM is calculated employing Geodesic Dilation, and then, using DTM and DSM we have extracted nDSM (H. Arefi, M. Hahn, 2005):

$$\begin{aligned} I &= DSM \\ J &= DSM - h \\ B &= StrEL \\ \delta_I^{(1)}(J) &= (J \oplus B) \wedge I \\ \delta_I^{(n)}(J) &= \delta_I^{(1)}(J) \circ \delta_I^{(2)}(J) \circ \dots \circ \delta_I^{(n)}(J) \\ nDSM &= DSM - \delta_I^{(n)}(J) \end{aligned} \quad (13)$$

Where  $StrEL$  = structure element  
 $\oplus$  = Morphological Dilation  
 $\wedge$  = minimum operation

**3.3.2 Third Section Result Layer:** At last using a threshold we would end up to the result layer:

$$nDSMRES = nDSM > P \quad (14)$$

Where  $P$  = threshold value

### 3.4 Final Processing

Now we have three result layers from three sections. In this part using a point-wise multiplication the total last result layer is computed:

$$Result = IMRES \times SLOPRES \times nDSMRES \quad (15)$$

Morphological operations are used to smooth the last result layer:

$$\begin{aligned} IMCLOSE &= closing(Result) \\ IMOPEN &= opening(IMCLOSE) \end{aligned} \quad (16)$$

Where  $closing$  = morphological closing  
 $opening$  = morphological opening

## 4. STUDY AREA

Our study area's data is gathered by ISPRS WG III/4. These data are captured over Vaihingen, Germany. ALS data has 4 points/m<sup>2</sup> point density and aerial images have 8cm ground pixel size and 11bits radiometric resolution. The field is divided into 3 areas. WG III/4 participants should test their methods for detection and extraction in these areas and submit their results to the committee to be evaluated. Each area could be seen fully or partially in some images. ALS point strips which cover three areas are shown in Fig. 4 (Franz Rottensteiner et al., 2011).

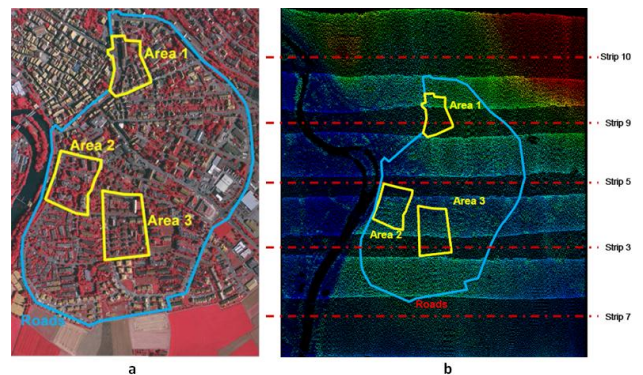


Figure . This figure shows test areas aerial images (a) and ALS points' strips (b).

## 5. PREPROCESSING INPUT DATA

In this paper we have worked on all of three areas. ALS strip bands 9, 5 and 3 are respectively used for Area 1, Area 2 and Area 3. First pulse range data from ALS points are used to calculate DSM. With knowing the fact that point density in ALS point cloud is 4 points/m<sup>2</sup>, employed grid intervals for DSM is 25cm. These steps are done using ENVI. As there is no Ground Control Points (GCPs) provided, for geo-referencing aerial images the DSM of each area is employed. To reduce processing process and make aerial images pixel size equal to DSM pixel size, aerial images are resampled to 25cm pixel size using Nearest Neighborhood (NN) algorithm. Then areas' aerial images and DSMs are cropped. These steps are done using Georeferencing panel in ArcGIS.

## 6. EXPERIMENTAL RESULTS

This section is performed using MATLAB. To be ensuring parameters are chosen correctly they are obtained for Area 1 and then used for the other two areas. Firstly NDVI and SI are calculated and then NDVI is divided on SI. This gave us ESI. Both NDVI and ESI indices are scaled and transferred to [0,255]. In continue after eliminating 2% of ESI and NDVI indices as outliers, they are reclassified into 25 classes by equal intervals. Now employing Eq. 6 and replacing X, A and B parameters respectively with 18, 0.3 and 0.7 LP layer is computed. Using 14 as a threshold in Eq. 7 the result layer in the first section is obtained. In the second section slope and RSC layers are computed using DSM and then employing 14 as a threshold in Eq. 12 the result layer for second section is computed. In the third section replacing h with 13m in Eq. 13 nDSM is calculated and then using 1m as a threshold in Eq. 14 the result layer in this section is obtained as well. Now we have three result layers from three sections. By multiplying these result layers the last result layer would be obtained. This layer is smoothed using morphological closing followed by morphological opening (Fig. 5). Structure element used in morphological operations is a disk with radius 1.

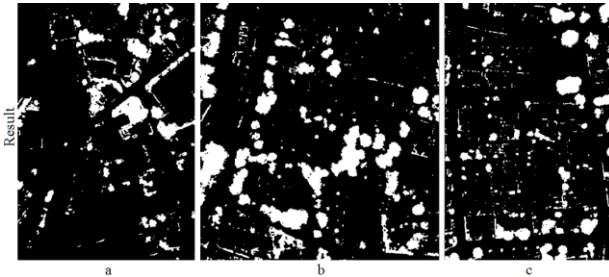


Figure 5. This figure shows the last result layer smoothed by morphological operations. White pixel show trees. a) Area1, b) Area2 and c) Area3

## 7. EVALUATION

Evaluation of the last results is performed by ISPRS WG III/4, so we are able to compare our results with other participants' results (Table 1). They used Fig. 6 as reference label images. This evaluation is performed on the basis of the method described in (Rutzinger et al., 2009). Our results details are reachable on the [ISPRS website](#). In the website our method is introduced as TAF. Considering Table 1 it is obvious our method is not the best method, but it is at the acceptable ranking. If we want to compare this method on the basis of boundaries RMS, this method in the Area 1 is in the second place by 1.5m and in both Area 2 and Area 3 it is in the first place by 1.4m and 1.2m values respectively. This is because completeness and correctness indices for this method are close enough, which means this method has good precision comparing to other methods. On the basis of completeness on both per-pixel and per-object level this method is in the third place in Area 1, in the sixth place in Area 2 and in the third place in Area 3. In the basis of correctness on a per-pixel level this method is in the fourth place in Area 1 and Area 2 and in the fifth place in Area 3. Fig. 7 represents the results in a graphical way. In this figure we can see large trees are detected very well, but there are some problems in detecting small trees. Because of using SI in this method trees which are in shadow areas are detected. Regarding to the results in Area 3 we can say this method has some problems in buildings boundaries.

Area1							
Participants	Compl area [%]	Corr area [%]	Compl obj [%]	Corr obj [%]	Compl obj 50 [%]	Corr obj 50 [%]	RMS [m]
CAL	37.2	80.1	30.5	53.9	50.0	100.0	1.4
HAN	41.4	69.2	27.6	46.1	50.0	64.7	1.4
HAN_J1	54.2	62.9	50.5	46.0	42.3	64.7	1.5
LJU	59.3	61.8	63.8	47.2	73.1	73.9	1.5
TUM	69.3	71.2	61.0	58.3	96.3	96.2	1.4
WHU	43.9	63.1	43.8	46.5	44.0	83.3	1.6
TAF	56.0	65.7	51.4	48.9	69.2	89.5	1.5
Area2							
Participants	Compl area [%]	Corr area [%]	Compl obj [%]	Corr obj [%]	Compl obj 50 [%]	Corr obj 50 [%]	RMS [m]
CAL	91.4	60.7	91.4	45.8	100.0	81.6	1.5
HAN	74.0	73.1	58.0	86.6	90.4	93.4	1.5
HAN_J1	81.5	64.1	74.1	63.2	95.2	78.8	1.5
LJU	88.9	59.2	79.0	55.2	98.8	80.2	1.5
TUM	72.0	78.5	63.0	82.4	89.3	98.6	1.4
WHU	64.2	71.5	48.8	70.9	75.6	93.5	1.5
TAF	68.5	69.5	54.3	63.1	77.1	97.0	1.4
Area3							
Participants	Compl area [%]	Corr area [%]	Compl obj [%]	Corr obj [%]	Compl obj 50 [%]	Corr obj 50 [%]	RMS [m]
CAL	83.8	58.6	81.3	28.1	100.0	78.6	1.4
HAN	55.9	77.0	29.0	68.9	74.2	100.0	1.3
HAN_J1	66.4	67.4	57.4	55.2	77.4	83.9	1.4
LJU	76.7	58.7	70.3	39.1	100.0	74.4	1.4
TUM	69.5	80.1	53.5	76.4	93.9	100.0	1.3
WHU	50.3	67.6	32.9	55.1	77.0	76.7	1.5
TAF	70.2	67.1	60.0	36.9	93.5	90.9	1.2

Table . Comparing results whit WG III/4 participants.

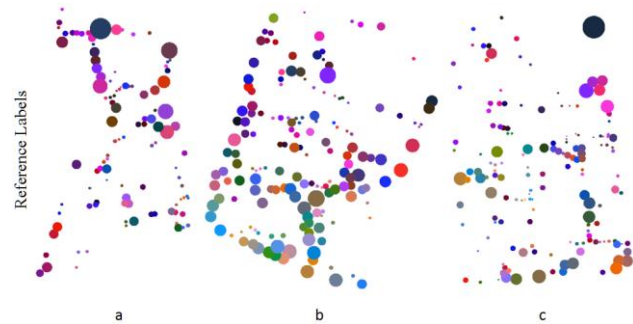


Figure 6. Reference labels used to evaluation. a) Area1, b) Area2 and c) Area3

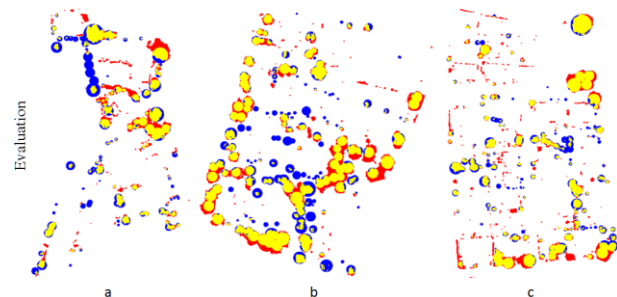


Figure 7. In this figure yellow pixels are True Positives, red pixels are False Negatives and blue pixels are False Positives.

## 8. CONCLUSION AND FUTURE WORK

Considering the results it is obvious indices used in this paper are good for the tree detection matter. This method is related to choosing precise values for some parameters, which are chosen manually. By changing these parameters the results could be improved, but it is a challenging job to select parameters manually. So Artificial Intelligence based methods, such as ANFIS, for determining parameters seem to be perfect. Replacing some parts of algorithm with alternatives could improve the results. For example, nDSM in this paper is calculated using Geodesic Dilation which it could calculate with other methods like Morphological Opening. Another example is method for calculating Shadow Index. We used just one method to calculate SI, which there are many methods to compute SI.

## 9. REFERENCES

Arevalo, V., Gonzalez, J., Valdes, J., & Ambrosio, G., 2006. "Detecting shadows in QUICKBIRD satellite images." ISPRS Commission VII Mid-term Symposium "Remote Sensing: From Pixels to Processes", (pp. 330-335). Enschede, the Netherlands.

Bulatov D., Solbrig P., Gross H., Wernerus P., Repasi E., Heipke C., 2011. "Context-based urban terrain reconstruction from UAV-videos for geoinformation applications." IAPRSIS XXXVIII-1-C22 (on CD-ROM)

Cai, D., Li, M., Bao, Z., Chen, Z., Wei, W., & Zhang, H., 2010. "Study on shadow detection method on high resolution remote sensing image based on HIS space transformation and NDVI index." 18th International Conference on Geoinformatics, 2010 (pp. 1-4). Beijing: IEEE.

Cheuk-Yan Wan\*, Bruce A. King, Zhilin Li, 2012. "An assessment of shadow enhanced urban remote sensing imagery of a complex city - Hong kong." International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XXXIX-B6, 2012 XXII ISPRS Congress, 25 August – 01 September 2012, Melbourne, Australia

Franz Rottensteiner, Caroline Baillard, Gunho Sohn, Markus Gerke, 2011: "ISPRS test project on urban classification and 3D building reconstruction."

Franz Rottensteiner and Simon Clode, 2008. "Topographic laser ranging and scanning Principles and Processing", Chapter 16. CRC Press

Grigillo, D., Kosmatin Fras, M., Petrovič, D. 2011. "Automatic extraction and building change detection from digital surface model and multispectral orthophoto." Geodetski vestnik 55, 1: 11–27.

H. Arefi, M. Hahn, 2005. "A morphological reconstruction algorithm for separating off-terrain points from terrain points in laser scanning data." ISPRS WG III/3, III/4, V/3 Workshop "Laser scanning 2005", Enschede, the Netherlands, September 12-14, 2005

Moussa, A., El-Sheimy, 2012. "A new object based method for automated extraction of urban objects from airborne sensors data." Accepted for publication in IAPRSIS XLIX-B3.

Niemeyer, J., Wegner, J.D., Mallet, C., Rottensteiner, F., Soergel, U., 2011. "Conditional random fields for urban scene

classification with full waveform LiDAR data." U. Stilla et al. (Eds.): PIA 2011, LNCS 6952, pp. 233-244.

Ono, Ak., Kajiwara, K., Honda, Y. and Ono, At. 2007. "Development of vegetation index using radiant spectra normalized by their arithmetic mean." V: Proceedings of the Asian Conference on Remote Sensing, Asian Association on Remote Sensing, Kuala Lumpur, Malezija, 12.-16. november 2007.

Rottensteiner, F., Sohn, G., Jung, J., Gerke, M., Baillard, C., Benitez, S., and Breitkopf, U, 2012. "The ISPRS benchmark on urban object classification and 3D building reconstruction." ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci., I-3, 293-298

Ritter, P. 1987. "A vector-based slope and aspect generation algorithm." Photogrammetric Engineering and Remote Sensing, 53:1109–1111.

Sarabandi, P., Yamazaki, F., Matsuoka, M., & Kiremidjian, A., 2004. "Shadow detection and radiometric restoration in satellite high resolution images." IEEE IGARSS '04., 6, pp. 3744-3747.

Tucker, C. J., 1979. "Red and photographic infrared linear combinations, for monitoring vegetation." Remote Sens. Environ., 8, 127–150

## 10. ACKNOWLEDGMENT

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> (in German).