SPECIAL ISSUE ON URBAN REMOTE SENSING

Orthorectified images from the Vaihingen dataset and normalized digital surface models of off-terrain-objects extracted.
For more information see "A Novel Building and Tree Detection Method From LiDAR Data and Aerial Images"
by Zarea and Mohammadzadeh, which begins on pp. 1869.

A Novel Building and Tree Detection Method From LiDAR Data and Aerial Images

Asghar Zarea and Ali Mohammadzadeh

Abstract—In recent decades, building and tree detection from LiDAR data and aerial imagery with high automation and accuracy level has been the focus of many researchers which was selected as the purpose of our research. At first, after data pre-processing, off-terrain objects (OTO) including trees and buildings were extracted from LiDAR data. Second, a number of features were produced as inputs of support vector machines (SVMs) to separate buildings from trees. In the SVM, an automatic procedure was used for selecting the training data. After separating the trees, mathematical morphology operations were used to eliminate small objects and fill small holes in the detected buildings and trees. Finally, k-means clustering algorithm was used to separate buildings with different heights. The obtained results for detected buildings and trees were evaluated by working group III/4 of ISPRS, which demonstrated a high rate of success. For completeness, correctness, and quality metrics in per area mode, average values of 88.70%, 95.60%, and 85.30% for buildings and 74.30%, 63.50%, and 52.10% for trees were obtained, respectively.

Index Terms—Building detection, LiDAR, mathematical morphology, support vector machines (SVMs), tree detection.

I. INTRODUCTION

FOR MANY countries, production of geospatial objects such as buildings and trees is one of the most important tasks. Main importance of building and tree detection is in updating of digital maps and GIS databases [1], [2], urban planning, environment protection, managing crises such as flood and earthquake, urban security, and real estate industry [3]. For this purpose, various works have been carried out which mostly are based on the following mathematical solutions: 1) active contour models [4]–[7]; 2) data classification and segmentation [8]–[14]; 3) Hough transformation [15]; 4) edge detection algorithms [3], [16]–[19]; and 5) mathematical morphology [20]–[22]. Among the mentioned solutions, classification-based algorithms have been widely employed and investigated. Khoshelham *et al.* [1] evaluated five different classification algorithms for building detection, which included Adaboost algorithm, maximum likelihood and minimum distance classifiers, Dempster-Shafer method, and thresholding a normalized digital surface model (nDSM). In this method, tree and other classes were obtained in classification process. In another method, Gerke and Xiao [13] detected building and tree classes by using segmentation and classification methods. Du [23] extracted buildings from high spatial resolution images using object-oriented classification. Classes of water, vegetation, and nonvegetation (including shadows, buildings, and roads) were hierarchically separated at two levels. Then, the buildings were separated from other classes at the third level. Miliaresis and Kokkas [10] applied segmentation and unsupervised object-based classification, in which manual operation was performed to identify buildings from the obtained results. Lach [24] used a pixel-based classification method to separate buildings from trees using manually selected training data (TD). Afterward, Lach proposed an automatic method for selecting TD in the presence of many single trees in the input data. Turlapaty *et al.* [25] separated tree and vegetation classes by using a threshold from normalized difference vegetation index (NDVI). Afterward, support vector machines (SVMs) classification was used to separate buildings from other remaining classes. The TD for their classification was selected by block classification method in a semiautomatic procedure which may contain some impure data. Thus, recently, there has been a tendency for increasing the automation of employed techniques that can be increased by the automatic labeling of building and tree classes or automatic selection of TD. Therefore, there is a need to investigate new approaches of increasing the automation level of TD selection in building and tree detection algorithms.

Another important problem in segmentation and classification of buildings and trees is their input features which would dramatically affect the results. For instance, Miliaresis and Kokkas used four features of mean and standard deviation of elevations and slopes for object-level classification. Lach used local texture features such as entropy, variance, Laplacian, and maximum difference that were derived from LiDAR data. Turlapaty *et al.* applied features such as mean, variance, skewness, kurtosis, energy, and two-dimensional (2-D) wavelet-based features of each block to separate buildings from nonbuildings. Mongus *et al.* [26] used three different geometric, textural, and regional features to detect buildings. To the best knowledge of the present authors, in the previous works, few features have been simultaneously used in building and tree detection and their performances have not been investigated. Thus, it is necessary to study the simultaneous use of appropriate features in a classification algorithm for building and tree detection.

In our proposed method, each submethod has its own strengths and weaknesses and its weaknesses are eliminated or

minimized by the subsequent one. For example, in the proposed method, separation of clung buildings was used to separate the building boundaries and minimized false positive pixels which have been created by the previous stage. The primary objective of this research was building and tree detection from LiDAR data and aerial image. In general, automatic selection of TD, examination of the effect of different features on building and tree detection, improvement of some algorithms or indices, and separation of clung buildings were the secondary aims of this study. According to the above introduction, we have six main contributions including: 1) appropriate data preprocessing; 2) OTOs detection; 3) various feature production and their simultaneous usage; 4) separating buildings and trees; 5) postprocessing and integration of classification results; and 6) clung buildings separation. The mentioned contributions are explained in the next sections in detail.

In this paper, the study area is described in Section II. The proposed methodology of building and tree detection is described in Section III. Section IV contains the implementation results of the proposed method and the discussions. Finally, conclusion is provided in Section V.

II. STUDY AREAS

In this study, the data provided by working group (WG) III/4 of ISPRS from Vaihingen area of Germany were used which included aerial images and LiDAR data. The aerial images were captured by a digital camera Intergraph/ZI DMC with the focal length of 120 mm at the elevation of 900 m above the ground on August 6, 2008.

The captured images over this area had 0.08-m spatial resolution and three bands of infrared (IR), red (R), and green (G). All the aerial images had interior- and exterior-orientation parameters. Airborne laser scanner data of the area were captured by a Leica ALS50 system on August 21, 2008 and included range and intensity of the first and last returns (LRs). The area was covered by 10 strips and mean strip overlap was 30%. The median point density in LiDAR data was 6.7 point per square meter. In the regions covered by one strip, the average point density was 4 point per square meter. A digital surface model (DSM) with 0.25-m spatial resolution was produced from original images by dense matching and was available for research. In the Vaihingen area, ISPRS-WGIII/4 has determined three study areas of area 1, area 2, and area 3. The study areas of this research were three study areas which were placed in the nadir of the aerial images. So, the amount of relief displacement error of images in three areas was low. The reference data for building and trees were prepared by ISPRS-WGIII/4. In this study, union of reference data for buildings and trees was used as the reference data for off-terrain objects (OTOs). The aerial image of study areas and reference data is shown in Fig. 1.

III. METHODOLOGY

Fig. 2 shows the proposed method for building and tree detection from LiDAR and aerial images. In this study, DSM was filtered based on a developed method to produce digital
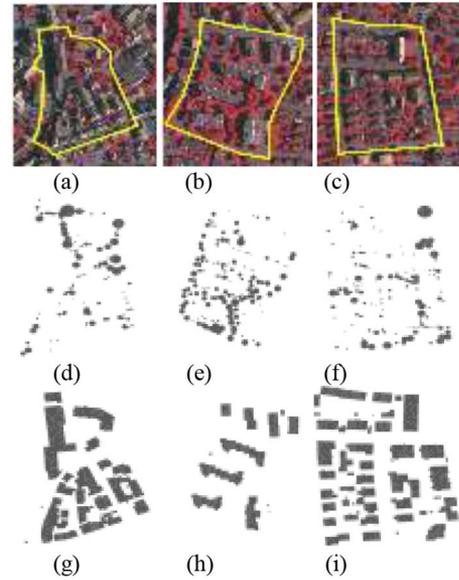
Fig. 1. (a)–(c) Aerial images of areas 1–3. Region of interest is specified with yellow lines in aerial images. (d)–(f) Reference data for trees of areas 1–3. (g)–(i) Reference data for buildings of areas 1–3.
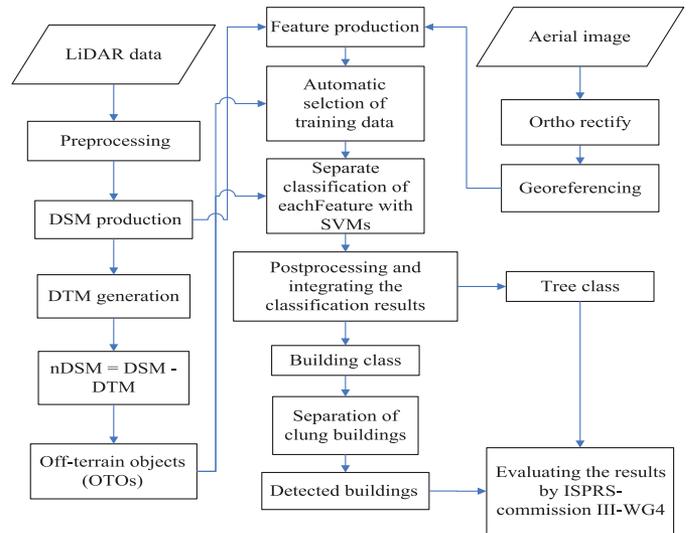
Fig. 2. Proposed methodology for building and tree detection from LiDAR data and an aerial image.

terrain model (DTM). So, nDSM was obtained by subtracting DTM from DSM. Then, OTOs were identified using a threshold from nDSM. As shown in Fig. 2, after the preprocessing step, DSM was fed to the feature production along with the georeferenced aerial ortho-image. Afterward, classification based on SVMs was used to separate buildings from trees in each feature.

TD for each feature classification was automatically selected. Finally, building and tree classes were separated after the postprocessing and integration of classification results. Also, clung buildings were separated by processing the building class. Commission III-WG4 of ISPRS evaluated the detected buildings and trees.

A. Data Preprocessing

In preprocessing as the first step, two important corrections were applied: 1) removal of a specific type of LiDAR noises and 2) ortho-rectification and georeferencing of input aerial images. In the correction type (1), LiDAR points that significantly had LR higher than their corresponding first return (FR) were considered to be noises. Thus, based on (1) and a threshold value of "Th1," this type of noise was removed

$$Point = \begin{cases} Noise, & \text{if } LR - FR > Th_1 \\ Pure, & \text{else.} \end{cases} \quad (1)$$

Detected noises were removed and replaced with an interpolated elevation from the neighboring points. For elevation interpolation, nearest neighbor (NN) method was used to keep the elevation leap along the edges of buildings [20].

In the correction type (2), input aerial images were ortho-rectified using DSM and interior- and exterior-orientation parameters. Collinearity model of (2) and (3) was used to remove relief displacements and ortho-rectification [27]

$$x_i = x_p - c$$
$$\times \frac{r_{11} \times (X_G - X_O) + r_{21} \times (Y_G - Y_O) + r_{31} \times (Z_G - Z_O)}{r_{13} \times (X_G - X_O) + r_{23} \times (Y_G - Y_O) + r_{33} \times (Z_G - Z_O)} \quad (2)$$

$$y_i = y_p - c$$
$$\times \frac{r_{12} \times (X_G - X_O) + r_{22} \times (Y_G - Y_O) + r_{32} \times (Z_G - Z_O)}{r_{13} \times (X_G - X_O) + r_{23} \times (Y_G - Y_O) + r_{33} \times (Z_G - Z_O)} \quad (3)$$

where (X_G, Y_G, Z_G), (X_O, Y_O, Z_O), $[r_{11}$ to $r_{33}]$, (x_i, y_i), and (x_p, y_p, c) are ground coordinates of object point, perspective center, elements of rotation matrix, image coordinates, and interior orientation parameters, respectively. Finally, ortho-rectified images were georeferenced using ground control points (GCPs) which were manually extracted from the input range images of LiDAR data.

B. OTOs Detection

Here, a bare earth detection algorithm was developed to identify OTOs based on scan labeling algorithm (SLA) proposed by Shan and Sampath [28]. SLA is an efficient filtering algorithm to obtain bare earth or DTM from LiDAR data. In this method, objects are considered to be between the discontinuities identified by height differences with neighboring points along the scan lines [28]. Thus, a continuity criteria (CC) is applied in both right–left direction (LRD) and left–right direction (RLD) for each scan line to label initial ground and nonground points with "1" and "0," respectively. Equation (4) shows the initial labeling function $\varphi(v_i)$ of this algorithm [29]

$$\varphi(v_i) = \begin{cases} 1, & \text{if } v_i^{LRD} + v_i^{RLD} = 2 \\ 0, & \text{else} \end{cases} \quad (4)$$

where v_i^{LRD} and v_i^{RLD} are initial labels of point v_i in LRD and RLD, respectively. There might be some nonground points
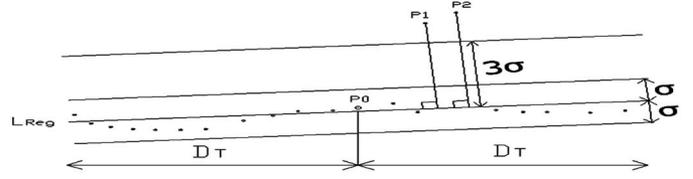
Fig. 3. Fitted 1-D regression to detect some possible nonground points.

which are initially labeled as ground points due to weak CC values [28]. So, using a one-dimensional (1-D) regression algorithm shown in (5), some of those nonground points would be separated from the ground points [28]

$$Z = a_0 + a_1 \times D, D < D_t \quad (5)$$

where D_t is neighbor radius of ground points and a_0 and a_1 are 1-D regression parameters. Considering Sigma (σ) as the standard deviation of point's distances from the regression line, any point with the distance of 3σ or higher would be considered the nonground. For example, if point P0 and line L_{reg} are considered as the investigated ground point and regression line, respectively; then points P1 and P2 would be identified as nonground points in Fig. 3.

Here, a developed SLA was proposed which had a number of advantages over the original SLA. In the proposed method, labeling happened in four RLD, LRD, up–down direction (UDD), and down–up direction (DUD), which led to more reliable identification of ground points than the original method. In the developed SLA, labeling function $\varphi_d(v_i)$ of point v_i was as follows:

$$\varphi_d(v_i) = \begin{cases} 1, & \text{if } v_i^{LRD} + v_i^{RLD} + v_i^{UDD} + v_i^{DUD} = 4 \\ 0, & \text{else} \end{cases} \quad (6)$$

where v_i^{LRD}, v_i^{RLD}, v_i^{UDD}, and v_i^{DUD} are temporary labels of point v_i in LRD, RLD, UDD, and DUD, respectively. Afterward, similar to the original SLA, linear regression method was used to eliminate some of the nonground points. Then, morphological opening was applied to the results to remove some nonground points that were not removed in the regression step. Finally, nDSM was attained by subtracting the resultant bare earth obtained by the developed SLA from DSM of the study area. According to (7), a height threshold "Th2" was used to identify OTOs binary image indicated by label "1" for detected buildings and trees

$$OTOs = \begin{cases} 1, & \text{if } nDSM > Th_2 \\ 0, & \text{else.} \end{cases} \quad (7)$$

Finally, another opening and closing operators were applied to remove very small OTOs and small gaps from OTOs.

C. Feature Production

In this section, a number of features are produced from DSMs and aerial images. These features would be inputs of the classification process for separating buildings from trees. For this

purpose, four features from aerial images including two different enriched vegetation indices (EVIs) and saturation and hue bands from IHS transformation, and five features from DSMs including gradient, variance, roughness, laplacian, and SSD [9] were produced. The above-mentioned features are described below.

1) Vegetation Indices Combined With Shadow Index: In this study, two slope-based vegetation indices (VIs) of NDVI and infrared-to-red ratio index (IRRI) [(8) and (9)] were used and developed to efficiently identify vegetation [30]–[32]. Vegetation had very high reflectance in IR band compared to R-band which resulted in high values for the mentioned indices

$$NDVI = (IR - R)/(IR + R) \qquad (8)$$
$$IRRI = IR/R. \qquad (9)$$

A problem was raised in shadow areas, in which vegetation had a low value for the mentioned indices. To solve this weakness, new indices were developed which enabled better detection of vegetation even in shadow areas. Grigillo *et al.* [33] used the difference between normalized R-band reflectance and original R-band reflectance as a shadow index (SI). Turlapaty *et al.* [25] used R and yellow band multiplication of a worldview-2 pan sharpened image as SI. Additionally, some shadow detection algorithms have been described by Shahtahmassebi *et al.* [34] and Chung *et al.* [35].

Our observations showed that shadows have low reflectance in R- and G-band of the aerial images. Equation (10) shows a new proposed SI in which low SI values indicated the shadows

$$SI = (G + R) \times G. \qquad (10)$$

By applying a threshold to the output of the SI, a binary image (BISA) which showed the shadow areas was produced. Also, a binary image of vegetation (BIveg) was calculated using a threshold from NDVI or IRRI bands. Then, the intersection of BISA with BIveg was demonstrated as a binary image (BISAV), which identified shadows over vegetation areas. Afterward, EVIs were obtained using the linear combination of VIs and BISAV according to (11). According to this enrichment, the shadow areas with vegetation would be strengthened and non-shadow areas would not be changed

$$EVIs = VIs + BI_{SAV} \times VIs/2. \qquad (11)$$

2) Saturation and Hue: Here, the aerial images with false color composition (G-R-IR) were transformed into an IHS color space to extract the desired hue and saturation features [36]. Intensity and saturation ranges were between "0" and "1"; and hue range was between 0 and 360. Zero value of hue represents red color and values of 120 and 240 show green and blue colors, respectively [36]. Saturation feature represents a dilution of color [36]. Vegetations in (G-R-IR) false color composition were shown in dark blue resulting in higher hue and saturations values than nonvegetation areas, which led to better discrimination.

3) Features Generated From Range Image of LiDAR Data: As previously mentioned, five features are generated from the range image of LiDAR data. In this research, to produce

gradient feature, a kernel with odd dimensions was placed on the range image. Then, the gradient images in horizontal (gradient$_h$), vertical (gradient$_v$), and diagonal (gradient$_d$) directions were produced and averaged according to (12)

$$Gradient = (gradient_h + gradient_v + gradient_d)/3. \qquad (12)$$

According to (13), Laplacian feature of each point was produced from the sum of the second derivatives of elevation with respect to x and y [36]

$$Laplacian = \partial^2 h/\partial x^2 + \partial^2 h/\partial y^2 \qquad (13)$$

where h(x, y) is the range image of LiDAR data and x and y represent the horizontal and vertical directions, respectively. If a kernel with the size of 3×3 pixels is considered, then second derivatives of elevation with respect to x and y can be calculated from (14) and (15) [36]

$$\partial^2 h/\partial x^2 = h(x + 1, y) - 2 \times h(x, y) + h(x - 1, y) \qquad (14)$$
$$\partial^2 h/\partial y^2 = h(x, y + 1) - 2 \times h(x, y) + h(x, y - 1). \qquad (15)$$

In the above equations, x and y are indices of the mentioned kernel in horizontal and vertical directions, respectively.

A kernel with odd dimensions was applied to the range image to produce SSD[1] [9], roughness, and variance features. SSD was obtained through fitted plan variance at all points within the kernel by least square method.

The features generated from the range images of LIDAR data were expected to have high values in elevated vegetation areas (trees) than building points, except for their edge points. Thus, in the postprocessing, morphological operators were used to remove those edge points. First, a threshold of "Th$_3$" was applied to feature image (F) to calculate a binary image (BI) as follows:

$$BI = \begin{cases} 1, & \text{if } F > Th_3 \\ 0, & \text{else.} \end{cases} \qquad (16)$$

PBI was obtained through applying a morphological closing followed by an opening on the BI to remove building boundary. Ultimately, the processed features (PF) were obtained according to (17)

$$PF = (1 - BI + PBI) \times F. \qquad (17)$$

Although the boundaries of buildings were removed, feature values of trees became weaker in the PBI image. Therefore, a maximum filter was applied to PF in order to compensate for the mentioned effect on trees.

D. Separating Buildings and Trees

Here, the produced features were overlapped with OTOs binary image to identify feature areas related to buildings and trees. Then, the identified feature areas were classified independently using SVMs to detect buildings and trees. SVM is a supervised classification method proposed by Cortes and

---

[1]Sum of squares of elevation deviation.

Vapnik [37]. To increase the automation level of the proposed algorithm, an approach was automatically developed to select the TD of SVM. The concept of automatic TD selection lies in the fact that feature values for trees are higher than the buildings due to essence of features. For example, values of NDVI and IRRI for trees are higher than buildings because of very high reflectance in IR-band compared to R-band for trees. This concept was declared in previous sections for each feature.

TD for each feature were independently selected. First, the produced features were overlapped with OTOs binary image to identify feature areas related to buildings and trees. Afterward, the identified area in each feature was independently classified into two classes using k-means clustering algorithm. Then, the classes were labeled as buildings and trees in an automatic procedure based on the fact that feature values for trees were higher than those for buildings. So, the classes that had high and low average feature values were labeled as trees and buildings, respectively. After clustering, (18) and (19) were used to select TD of each feature for building and tree classes, respectively

$$TD_{Build} = \begin{cases} 1, & \text{if } FV \langle m_1 + d \times s_1 \text{ and } FV \rangle m_1 - d \times s_1 \\ 0, & \text{else} \end{cases}$$
(18)

$$TD_{Trees} = \begin{cases} 1, & \text{if } FV \langle m_2 + d \times s_2 \text{ and } FV \rangle m_2 - d \times s_2 \\ 0, & \text{else} \end{cases}$$
(19)

where FV is feature value, (m1, s1) and (m2, s2) are center and standard deviation of building and tree classes in the related features, respectively. "d" is constant coefficient in (0,3], and labels "1" and "0" represent the pixel which is selected as TD and non-TD sets, respectively.

In the classes, by decreasing the value of parameter "d," more reliable TD points were selected; but, their percentage was dramatically reduced and vice versa. Therefore, an appropriate value for parameter "d" was arbitrarily determined by users. Finally, a definite number of selected TD points were randomly chosen as the final selected TD points.

E. Postprocessing and Integration of Classification Results

In postprocessing, morphological opening and closing operators were applied to remove small objects and fill holes of buildings and trees in each SVM classification result. Afterward, nine postprocessed binary images were obtained for each class. It is necessary to note that buildings and trees are illustrated in independent binary images which values of 1 and 0 refers to buildings/trees and nonbuildings/nontrees, respectively. These nine binary images for building/trees (a total of 18 images) are added together to produce integrated results for buildings/trees. According to (20), a pixel is assigned to building/tree class if the total amount of the pixel in the nine binary images of building/tree is greater than threshold "Th_4"

$$\text{Pixel}^i = \begin{cases} \text{Building}, & \text{if } SL_B^i \geq Th_4 \\ \text{Tree}, & \text{if } SL_T^i \geq Th_4 \\ 0, & \text{else} \end{cases}$$
(20)

where SL_B^i and SL_T^i are the sum of labels for a pixel in the nine processed binary images of building and tree, respectively. Also, "Th_4" is a constant integer value within [1], [9] that controls the entirety and validity of detected buildings and trees. In other words, (20) expresses that pixels are assigned to the building/tree class if they are classified as building/tree pixels at least in a certain percentage of processed binary images.

F. Separation of Clung Buildings

In this section, clung buildings with elevation difference of above 1.5 m are separated as much as possible. First, different labels were assigned to the detected buildings by connected component analysis (CCA) algorithm, except for clung buildings which take the same label. For separating clung buildings, a k-means clustering algorithm with two classes was independently applied to the nDSM of each detected building. Afterward, mathematical morphological operations of opening followed by closing were applied to remove small clusters and fill the holes, which would increase the validity of the detected buildings, but would decrease their entirety. Buildings with two main separated clusters were considered clung buildings and unique labels were given to each separated building. If one of the clusters of a building had low population percentage of the building, then that cluster would be separated from the building with a given unique label, but would not be considered a clung building.

G. Evaluation Method

Reference data provided by ISPRS-WGIII/4 were used for the evaluation of the previous steps, except for the final detection results which were carried out by ISPRS-WGIII/4 itself. To evaluate the results, completeness (CP), correctness (CR), and quality (Q) metrics were used in per area and per object modes based on the method described by Rutzinger et al. [38]. In addition to the mentioned metrics, average root-mean-square (rms) distances between the corresponding vertices of reference and detected building polygons were used as geometrical accuracy for the evaluation of the final detected buildings and trees result. Details of the evaluation based on the rms metric were described in http://www2.isprs.org/tl_files/isprs/wg34/docs/EvaluationObjectDetection.pdf

According to (21), overall accuracy of OTOs classification was used to interpretation some parts of the results [1]

$$OA = (TP + TN) / (TP + TN + FP + FN)$$
(21)

where TP and TN are the number of pixels which classified correctly as building and tree, respectively, FP and FN are the number of tree and building pixels which classified as building and tree, respectively.

IV. RESULTS AND DISCUSSION

A. Preprocessing

Considering the average elevation accuracy of 0.15 m for LiDAR data [39], [40] height difference accuracy (σ_{total}) of

Thisarticlehasbeenacceptedforinclusioninafutureissueofthisjournal.Contentisfinalaspresented,withtheexceptionofpagination.

6   IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING
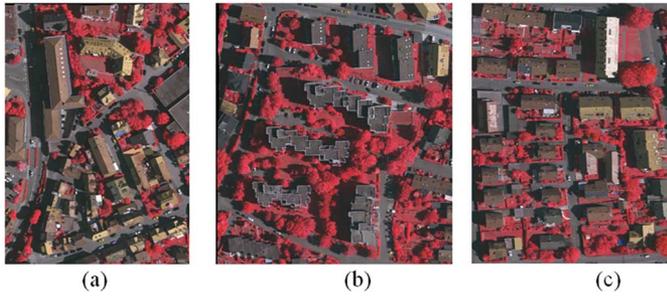
Fig. 4. (a)–(c) Ortho-rectified images of areas 1–3.

0.21 m was calculated based on error propagation between the FRs and LRs [refer to (22)]

$$\sigma_{total} = \sqrt{\sigma_{FR}^2 + \sigma_{LR}^2} = \sqrt{0.15^2 + 0.15^2} = 0.21 \text{ m} \quad (22)$$

where σFR and σLR are the FR and LR elevation accuracies, respectively. A user-defined value of 0.3 m was chosen for "Th1" which was larger than the height difference accuracy. By applying (1), about 0.14%, 0.56%, and 0.2% of total LiDAR points were detected as noise in areas 1–3, respectively. Afterward, grid-based LiDAR range data were produced with the resolution of 0.25 m, in which ISPRS-WGIII/4 used the same resolution as their DSM products. Consequently, ortho-rectified image with 0.25-m resolution was obtained using range data and aerial image, as shown in Fig. 4. In this research, 73, 84, and 71 GCPs with uniform distribution were used in the georeferencing of the ortho-rectified images of areas 1–3, respectively. The GCPs were selected by visual matching between ortho-rectified image and DSM of the study areas. Rotation, transition, and scale (RST) method were used to geo-reference the ortho-rectified images. Average RMSE of 0.9, 0.97, and 0.95 DSM pixel size are calculated for 73, 84, and 71 GCPs of areas 1–3, respectively. Also, RMSE of 0.96, 1.08, and 1.01 DSM pixel size are calculated for 10 ground check points of areas 1–3, respectively.

B. OTOs Detection

In the proposed method, gradient between adjacent pixels was used as the continuity criterion (CC). In the original and developed SLA, allowable slope value to comply with CC was considered 60% in all study areas. For both algorithms, a kernel with the size of 7×7 pixels was used in all study areas, instead of the neighborhood radius mentioned in (5). Then, in the developed SLA, a kernel with the size of 9×9 pixels was used as the structuring element (SE) of morphological opening to remove nonground points. After the final diagnosis of terrain points, a cubic interpolation followed by an NN one was applied to the results of the original and developed SLA to produce DTM.

Fig. 5 shows the efficiency of the developed SLA in DTM and nDSM of OTOs production compared to the results obtained by applying the original SLA. A threshold (Th$_2$) of 1.5 m was used to remove low height objects from nDSM to produce nDSM of OTOs. Afterward, objects with planimetric dimensions of below 2.5 m × 2.5 m were removed from OTOs binary images (orthogonal projection of nDSM of OTOs on XY
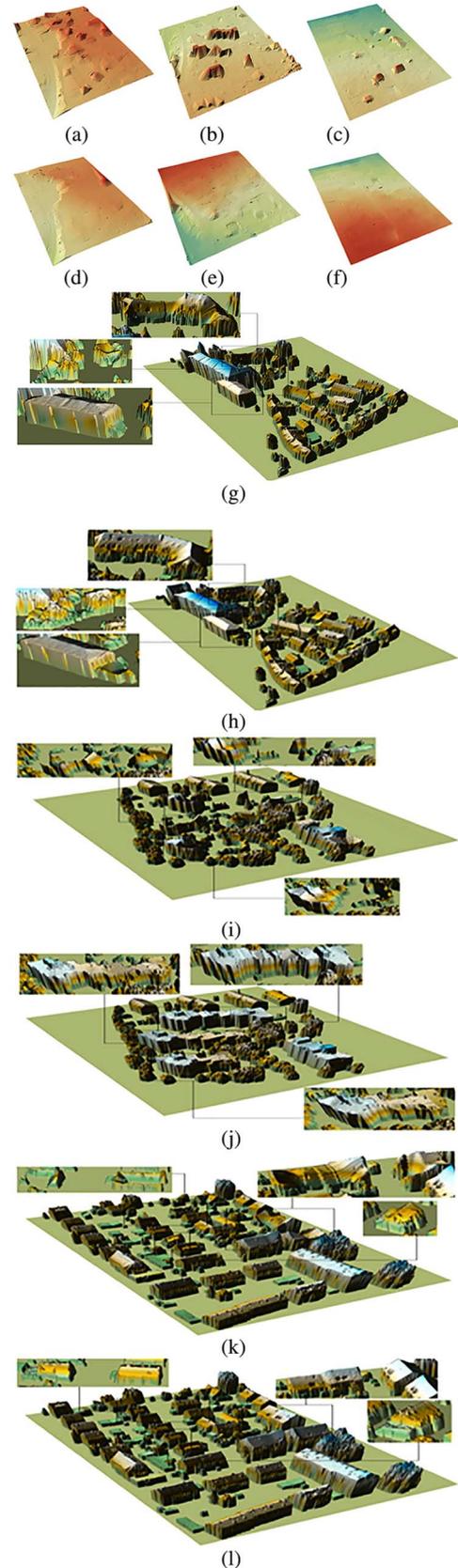
Fig. 5. (a)–(c) Generated DTM by original SLA in areas 1–3, respectively. (d)–(f) Generated DTM by developed SLA in areas 1–3, respectively. (g), (i), and (k) Produced nDSM of OTOs based on original SLA in areas 1–3, respectively. (h), (j), and (l) Produced nDSM of OTOs based on developed SLA in areas 1–3, respectively.

Thisarticlehasbeenacceptedforinclusioninafutureissueofthisjournal.Contentisfinalaspresented,withtheexceptionofpagination.

ZAREA AND MOHAMMADZADEH: NOVEL BUILDING AND TREE DETECTION METHOD FROM LiDAR DATA AND AERIAL IMAGES
7

TABLE I
ASSESSMENTS RESULTS OF DETECTED OTOs BASED ON ORIGINAL AND DEVELOPED SLA IN PER AREA MODE FOR CP, CR, AND Q

Method		Metric	Area 1	Area 2	Area 3
Original SLA	Before processing by morphological operators	CP (%)	85.60	83.56	90.12
		CR (%)	86.60	72.67	84.78
		Q (%)	75.51	63.58	77.57
	After processing by morphological operators	CP (%)	88.87	83.29	84.07
		CR (%)	84.52	73.24	91.83
		Q (%)	76.36	63.85	78.22
Developed SLA	Before processing by morphological operators	CP (%)	89.29	89.59	92.87
		CR (%)	84.21	73.08	84.19
		Q (%)	76.49	67.36	79.08
	After processing by morphological operators	CP (%)	88.53	81.95	86.76
		CR (%)	85.20	81.13	91.43
		Q (%)	76.73	68.83	80.24

plane) by a morphological opening with the SE size of 9×9 pixels. Afterward, small holes with the planimetric dimensions of below $1.5 \text{ m} \times 1.5 \text{ m}$ were filled by applying a morphological closing with the SE size of 5×5 pixels. The final output for OTOs binary images was compared with the reference data of OTOs which were obtained with union of reference data for buildings and trees. The results of the corresponding accuracy assessment are presented in Table I. According to Table I, it can be easily seen that the Q metric of the developed SLA in areas 1–3 in comparison to the original SLA for both before/after morphology steps had improved accuracies by about 0.7%, 4.4%, and 2%, respectively. Additionally, applying morphological step improved the Q metric on average by about 1% for the developed SLA.

C. Feature Production and Postprocessing

The produced features related to LiDAR DSM were postprocessed. In all study areas, value of parameter "Th3" mentioned in (16) was considered 0.1 for SSD, variance, and roughness features and 0.3 for gradient and Laplacian features. Then, a morphological closing with the SE of 5×5 pixels followed by an opening operator with the SE of 19×19 pixels was applied to remove building edges.

Finally, a maximum filter with the size of 3×3 pixels was applied to compensate effects of trees mentioned in Section III-C. Fig. 6 shows the efficiency of EVIs in shadow areas in comparison to VIs. As can be seen in this figure, EVIs have high value for vegetation in shadow areas in comparison to VIs.

D. Separating Buildings and Trees

As mentioned in Section III-D, parameter "d" should be appropriately chosen to select a sufficient number of TD in SVMs classification. Fig. 7 shows the diagram of variation of average overall accuracy of nine classifications for the change of parameter "d" mentioned in (18) and (19). For study areas 1, 2, and 3, the standard deviation values of 0.2180, 0.0192, and
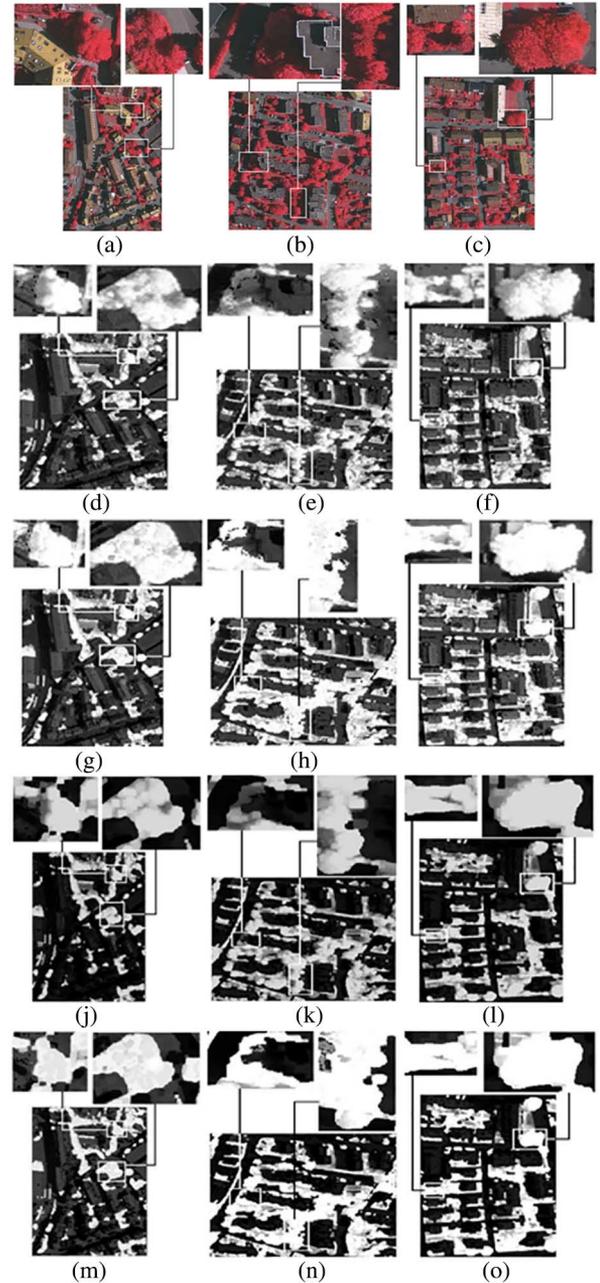
Fig. 6. (a)–(c) Ortho-rectified images of areas 1–3: in the white boxes, shadow overlaid on whole or some part of trees. (d)–(f) Traditional NDVI of areas 1–3. (g)–(i) Enriched NDVI of areas 1–3. (j)–(l) Traditional IRRI of areas 1–3. (m)–(o) Enriched IRRI of areas 1–3.

0.0034 were observed in the overall accuracy of obtained classification, respectively. According to these experiments, it would be deduced that the classification results were not very sensitive to small variations in parameter "d." By choosing value 1 for parameter "d," possible TD set would be identified, out of which only 50 random samples are selected as TD for each of building and tree classes in SVMs. Here, a Gaussian radial basis (RBF) function kernel was used in the SVMs classification to achieve better results [41].

In this section, detected buildings and trees for traditional NDVI and IRRI are presented to compare enriched NDVI (ENDVI) and IRRI (EIRRI); but, they do not contribute in
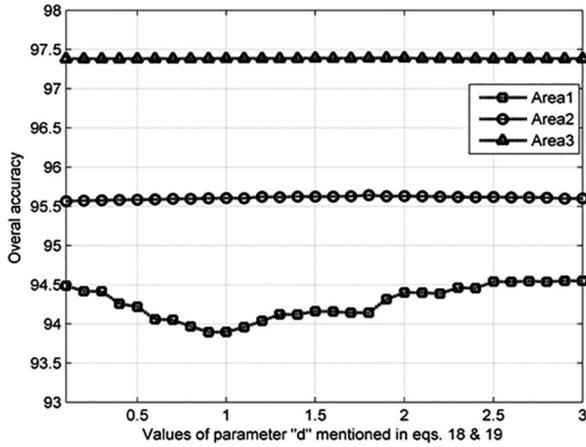
Fig. 7. Average overall accuracy of classifications as a function of parameter "d."

TABLE II
EVALUATING DETECTED BUILDINGS FROM EACH FEATURE
IN PER AREA MODE

Input features		Area 1			Area 2			Area 3		
		CP	CR	Q	CP	CR	Q	CP	CR	Q
Aerial image-related features	NDVI	90.1	90.6	82.4	87.1	85.7	76.2	87.4	92.3	81.5
	IRRI	90.2	91.2	83.0	87.0	86.7	76.8	87.9	92.3	81.9
	ENDVI	**90.4**	91.8	83.7	**87.3**	**88.6**	**78.5**	88.3	94.0	83.5
	EIRRI	90.3	92.4	**84.1**	87.0	87.5	77.4	88.2	**94.5**	**83.9**
	Saturat	71.5	89.5	66.1	84.5	80.3	70.0	89.2	91.6	82.5
	Hue	77.0	92.8	72.7	83.2	81.6	70.1	87.0	93.2	81.8
DSM-related features	Grad	87.9	89.7	79.9	85.7	86.7	75.8	85.8	92.5	80.2
	Lap	89.6	88.0	79.9	86.1	86.4	75.8	87.3	91.1	80.4
	SSD	90.8	89.3	81.9	86.9	82.0	72.9	88.6	91.8	82.1
	Rough	68.3	**98.2**	67.5	87.0	77.4	69.4	89.6	83.8	76.4
	Var	73.5	92.8	69.5	86.4	82.0	72.7	**88.9**	85.7	77.5

The best values per column are highlighted by bold font.

TABLE III
EVALUATING DETECTED TREES FROM EACH FEATURE
IN PER AREA MODE

Input features		Area 1			Area 2			Area 3		
		CP	CR	Q	CP	CR	Q	CP	CR	Q
Aerial image-related features	NDVI	49.9	63.0	38.7	67.0	68.8	51.4	57.4	**82.0**	51.0
	IRRI	52.4	61.1	39.3	68.7	67.1	51.4	59.4	79.0	51.3
	ENDVI	53.3	67.0	42.2	68.8	69.2	**52.7**	60.5	80.2	52.6
	EIRRI	**55.5**	64.3	**42.4**	**70.1**	67.3	52.3	**62.3**	78.5	**53.2**
	Saturat	42.2	54.9	31.3	54.5	**76.4**	46.7	55.8	79.7	48.9
	Hue	43.5	65.3	35.4	62.8	61.4	45.0	57.7	74.8	48.3
DSM-related features	Grad	46.5	60.6	35.7	65.2	72.0	52.0	53.0	74.8	45.0
	Lap	47.3	63.4	37.1	65.9	69.9	51.3	54.2	77.1	46.7
	SSD	49.0	**67.8**	39.7	67.2	71.0	52.7	52.1	80.2	46.2
	Rough	43.5	57.7	33.0	60.2	70.3	48.0	33.4	80.3	30.9
	Var	44.7	49.0	30.5	59.8	70.1	47.6	36.5	**83.2**	34.0

The best values per column are highlighted by bold font.
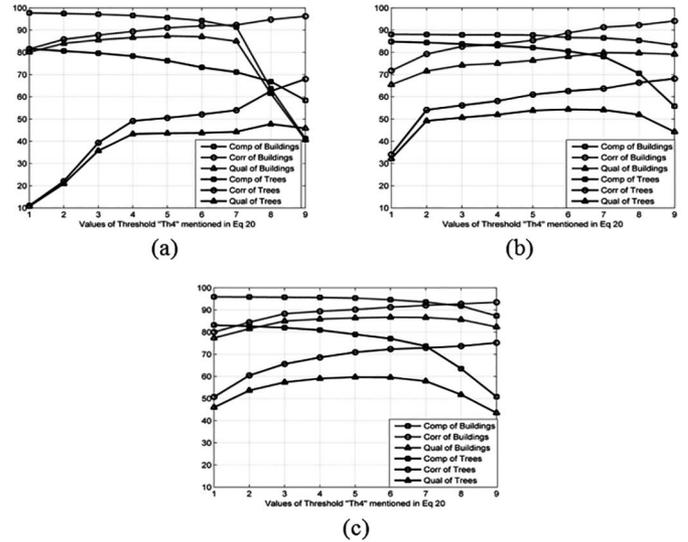
Fig. 8. Effect of parameter "Th4" in CP, CR, and Q of detected buildings and trees. (a) Area 1. (b) Area 2. (c) Area 3.

separating the building from trees. To improve the detected buildings and trees, a morphological opening followed by a closing operator with the SE size of 7×7 pixels was applied to remove small objects and fill small holes in each detected building and tree results. Tables II and III show the evaluation of results based on the reference data provided by WGIII/4 of ISPRS under "urban classification and three-dimensional (3-D) building reconstruction" test project in 2013. According to Tables II and III, ENDVI and EIRRI had the best results compared to the ones obtained by the remaining features. Additionally, average qualities of detected buildings and trees in three study areas were increased by almost 2% and 1.6% in enriched NDVI and IRRI compared to traditional NDVI and IRRI, respectively.

E. Integration Classification Results

For combining classification results, threshold "Th4" was considered to be 7, i.e., the integration result for a pixel of interest would be building/tree if at least 75% or 7/9 of the classifications results were labeled as building/tree. According to Fig. 8, selecting such a high value for parameter "Th4" would increase the CR of detected buildings or trees and decrease their CP and vice versa. Fig. 9 shows detected buildings and trees from the integration of the classification results and Table IV represents their evaluation results based on pixel-level indices. By comparing Table IV with Tables II and III, it can be deduced

Thisarticlehasbeenacceptedforinclusioninafutureissueofthisjournal.Contentisfinalaspresented,withtheexceptionofpagination.

ZAREA AND MOHAMMADZADEH: NOVEL BUILDING AND TREE DETECTION METHOD FROM LiDAR DATA AND AERIAL IMAGES 9
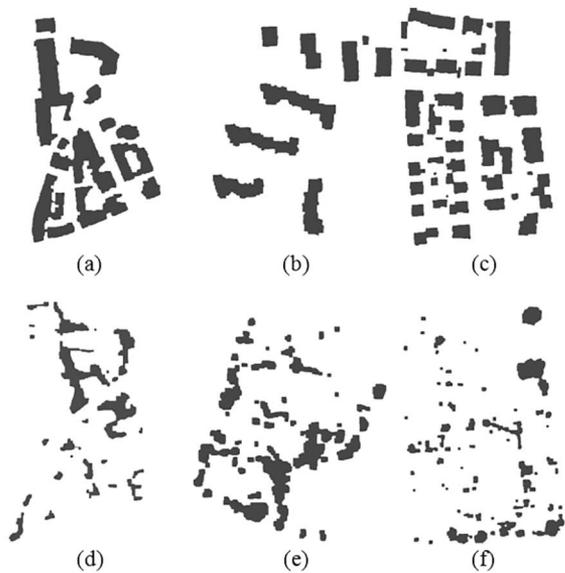
Fig. 9. Results of building and tree detection after combining the classification results. (a) and (d) Area 1. (b) and (e) Area 2. (c) and (f) Area 3.

TABLE IV
EVALUATION RESULTS OF DETECTED BUILDINGS AND TREES
IN PER AREA MODE

Detection type	Study area	CP (%)	CR (%)	Q (%)
Buildings	Area 1	91.4	**92.4**	85.0
	Area 2	86.5	91.3	79.9
	Area 3	**93.6**	92.1	**86.6**
	Average	90.5	91.9	83.8
Trees	Area 1	71.1	54.0	44.3
	Area 2	**78.1**	63.7	54.1
	Area 3	73.7	**72.9**	**57.9**
	Average	74.3	63.5	52.1

that two metrics of CP and Q increased after the integration of classification results.

F. Separating Clung Buildings With Different Elevations

Fig. 10 illustrates the separated clung buildings based on k-means algorithm with successful results. As can be seen in this figure, individual buildings with different roof elevations were not divided and just clung buildings were discriminated. At this stage, a morphological opening with the SE size of 9×9 pixels followed by a closing operator with the SE size of 5×5 pixels was applied to each class to improve the results. By comparing Tables IV and V, it can be express that metrics of CR and Q increased after separating clung buildings by about 3.5% and 1.5%, respectively.

G. Discussion

A brief evaluation of the obtained results was published with abbreviations KNTU/KNTU_mod for tree/building in ISPRS link at http://www2.isprs.org/commissions/comm3/
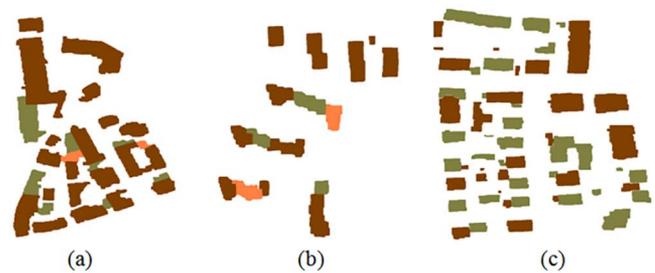
Fig. 10. Detected buildings after the separation of clung buildings. (a) Area 1. (b) Area 2. (c) Area 3.

TABLE V
EVALUATION RESULTS OF DETECTED BUILDINGS IN PER AREA MODE
AFTER SEPARATING CLUNG BUILDINGS

Detection type	Study area	CP (%)	CR (%)	Q (%)
Buildings	Area 1	**91.4**	94.3	86.6
	Area 2	86.5	93.6	81.7
	Area 3	88.3	**99.0**	**87.5**
	Average	88.7	95.6	85.3

TABLE VI
ISPRS-WGIII/4 AVERAGE EVALUATION RESULTS FOR DETECTED
BUILDINGS OF THE PROPOSED METHOD COMPARED TO OTHER
METHODS [42] (ABOVE-MENTIONED LINK)

ID	Per area			Per object			Per object > 50 m²		
	CP	CR	Q	CP	CR	Q	CP	CR	Q
KNTU_Mod	88.7	95.6	85.3	82.7	99.3	82.2	**100**	**100**	**100**
UMT	92.3	87.5	81.5	80.0	98.6	79.1	99.1	100	99.1
UMT	92.4	86.0	80.3	80.9	95.8	78.1	98.8	97.2	96.0
MON	92.7	88.7	82.8	82.7	93.1	77.7	99.1	**100**	99.1
VSK	85.8	**98.4**	84.6	79.7	**100**	79.7	97.9	**100**	97.9
WHU	87.3	91.6	80.8	77.6	98.1	76.5	97.4	97.9	95.4
WHU	89.7	90.9	82.3	83.0	97.5	81.3	99.1	98.0	97.2
HAN	91.5	92.5	85.2	81.5	72.7	62.4	**100**	95.8	95.8
HAN	90.2	93.2	84.6	85.1	69.6	61.9	**100**	**100**	**100**
MAR	87.0	97.1	84.8	78.2	96.2	75.7	99.1	100	99.1
MAR	89.7	95.2	85.8	80.6	93.7	76.5	99.1	98.9	98.0
TON	77.7	97.7	76.3	67.5	98.9	66.9	92.7	98.8	91.6
DLR	93.3	96.0	**89.8**	80.3	99.0	79.6	**100**	**100**	**100**
FIE	89.0	86.9	78.5	78.6	**100**	78.6	**100**	**100**	**100**
HAN	93.6	90.3	85.0	80.3	88.8	73.0	97.4	97.2	94.6
RMA	92.8	90.2	84.2	82.7	81.0	68.1	**100**	**100**	**100**
MEL	88.0	79.2	71.4	75.9	76.1	59.7	97.4	81.3	78.8
CAL1	89.8	95.1	85.8	76.2	100	76.2	96.5	**100**	96.5
CAL2	89.2	97.2	87.0	78.2	**100**	78.2	**100**	**100**	**100**
LJU1	94.2	94.6	89.4	83.0	**100**	83.0	**100**	**100**	**100**
LJU2	**94.6**	94.4	89.5	**87.9**	**100**	**87.9**	**100**	**100**	**100**
TEH	76.7	93.8	73.0	75.7	90.3	70.0	85.7	**100**	85.7
TUM	89.7	92.9	83.9	80.9	99.0	80.2	99.1	**100**	99.1
WHU	80.3	89.5	73.2	66.6	55.0	42.0	83.6	95.7	80.7
ZJU	92.8	96.4	89.7	76.4	97.0	74.8	99.1	**100**	99.1
ITCM	92.7	80.9	75.9	84.8	51.2	47.1	99.1	88.9	88.0
ITCR	91.4	90.6	83.5	80.0	70.6	60.0	98.2	**100**	98.2

The best values per column are highlighted by bold font.

wg4/results.html in comparison to the results obtained by other researchers [42]. It should be noted that the buildings with the area size of below 2.5 m² were not included in the evaluation process. It is necessary to express that the results mentioned in [42] were the average results of areas 1–3, whereas those mentioned in the above link were the results of areas 1–3. By

TABLE VII

ISPRS-WGIII/4 AVERAGE EVALUATION RESULTS FOR DETECTED TREES OF THE PROPOSED METHOD COMPARED TO OTHER METHODS [42]

ID	Per area			Per object			Per object > 50 m^2		
	CP	CR	Q	CP	CR	Q	CP	CR	Q
KNTU	74.3	63.5	52.1	66.5	68.0	**49.9**	92.9	74.5	70.5
HANC1	57.1	73.1	47.0	38.2	67.2	33.2	71.5	86.0	66.2
HANC2	67.4	64.8	49.1	60.7	54.8	40.9	71.6	75.8	59.2
DLR	58.8	**76.6**	49.8	47.2	**75.1**	41.0	70.0	87.1	63.2
CAL1	70.8	66.5	48.0	67.7	42.6	31.5	83.3	86.7	70.1
CAL2	67.2	70.4	52.8	57.2	75.0	48.4	82.6	82.9	73.3
LJU1	**75.0**	59.9	49.4	**71.0**	47.2	39.6	90.6	76.2	70.6
LJU2	63.8	67.2	48.3	43.7	66.4	36.2	81.3	86.9	72.8
TEH	56.9	51.3	36.9	50.5	20.4	17.0	80.8	52.0	46.3
TUM	70.3	**76.6**	**57.8**	59.2	72.4	48.0	**93.2**	**98.3**	**91.6**
WHUZ	52.8	67.4	42.2	41.8	57.5	31.9	63.5	84.5	56.9
ITCM	49.2	69.4	40.4	37.4	65.3	31.2	72.5	87.9	65.8
ITCR	64.0	66.9	47.9	51.2	65.4	41.0	85.8	87.5	75.4

Best values per column are highlighted by bold font.

TABLE VIII

AVERAGE RMS VALUES FOR DETECTED BUILDINGS OF THE PROPOSED METHOD COMPARED TO OTHER METHODS [42], (ABOVE-MENTIONED LINK)

ID	RMS	ID	RMS	ID	RMS
KNTU_Mod	0.83	MAR1	0.83	CAL2	0.77
UMTA	0.87	MAR2	0.83	LJU1	0.73
UMTP	0.97	TON	0.90	LJU2	0.75
MON	0.93	DLR	0.73	TEH	1.00
VSK	0.87	FIE	1.20	TUM	1.03
WHUY1	0.83	HAND	0.83	WHUZ	1.10
WHUY2	0.90	RMA	0.90	ZJU	0.63
HANC1	**0.67**	MEL	1.10	ITCM	1.13
HANC2	0.83	CAL1	0.73	ITCR	0.93

TABLE IX

AVERAGE RMS VALUES FOR DETECTED TREES OF THE PROPOSED METHOD COMPARED TO OTHER METHODS [42]

ID	RMS	ID	RMS	ID	RMS
KNTU	1.50	CAL2	1.33	WHUZ	1.53
HANC1	1.40	LJU1	1.47	ITCM	1.50
HANC2	1.47	LJU2	1.43	ITCR	1.53
DLR	**1.30**	TEH	1.60		
CAL1	1.43	TUM	1.37		

comparing the average results, it can be seen that detected buildings and trees have third and first Q metric in per object mode, respectively. It should be expressed that Q metric was not specified in the above-mentioned link but can be calculated from CP and CR metrics [38].

Moreover, the average evaluation results of proposed method for building detection achieved maximum CP, CR, and Q metrics in mode of per object larger than 50 m^2 among other mentioned methods.

Tables VI and VII demonstrate the obtained accuracies per area, per object, and per object larger than 50 m^2 for the proposed method compared to other methods [42] (above-mentioned link). According to Tables VIII and IX, average rms of differences between the corresponding vertices of detected building/tree and their reference data was approximately 0.83 and 1.50 m, respectively. Once more, it can be easily deduced that the obtained RMSE by the proposed method for building

detection was in the top 25% compared to other RMSEs which were between 0.6 and 1.2 m. The RMSE's of building and trees for other works are in domains of (0.67–1.20) and (1.3–1.6), respectively. Our achieved RMSE's for buildings and trees are 0.83 and 1.5, respectively, which indicates acceptability of the results.

One of the important errors in building detection is the existence of nonbuilding pixels (type II error[2]) which was dramatically reduced in the proposed method owing to the following reasons. As noted earlier, low values of "Th_4" mentioned in (20) reduced CR and increased CP and type II error of the detected buildings and vice versa. Thus, a high value was chosen for $Th_4 = 7$. Also for separation of clung buildings from other remaining ones, especially individual buildings, nonbuilding pixels were removed by morphological operators, which resulted in high CR accuracy.

V. CONCLUSION

The purpose of this study was to present a method which identified the buildings and trees from LiDAR and aerial images. The EVIs production, automatic TD selection, and clung building separation were initiatives of the proposed method. Also, SLA was developed in this study to efficiently detect OTOs. Q metric of OTOs detection based on the developed SLA demonstrated an average improvement by about 2.3% in comparison to the original SLA. In this study, an independent classification was used for each feature to separate buildings from trees and finally classification results were integrated. The procedure proposed for automatic TD selection was successfully determined to be correct and sufficient TD for SVMs classification. In addition, average qualities of SVMs classification based on ENDVI and EIRRI increased by almost 2.0% and 1.6% in comparison to the traditional NDVI and IRRI, respectively. It can be expressed that the uses of morphological operators slightly improved results in each section. In average, morphological operators increased the Q metric of OTOs detection by about 1%. In the separation of clung building with different elevations, nonbuilding pixels were filtered or reduced as much as possible using morphological operators, which resulted in average 1.5% increase in the Qmetric of the detected buildings. A high value for the CR of building detection result played an important role in 3-D building reconstruction. Comparing to type I[3], type II error more severely affects the determination of 3-D planes or roof structure of buildings from LiDAR data and would result in some difficulties. Evaluation results demonstrated the efficiency of the proposed method; in per area mode, average CP, CR, and Q of 88.7%, 95.6%, 85.3% and 74.3%, 63.5%, 52.1% were obtained for building and tree detection, respectively. In per object mode, the proposed method achieved high and maximum Q for detected buildings and trees, respectively; in comparison to other mentioned methods and had still high values for CP and CR metrics. In addition, average geometric accuracies of the building and tree boundaries were approximately 0.83 and 1.5 m, respectively; that was also convincing in comparison to

---

[2]Nonbuilding pixels that are classified as building pixels.

[3]Building pixels that are classified as nonbuilding pixels.

the accuracy of others method. In general, the proposed method can be used as a successful method for the detection of buildings and trees with different sizes using LiDAR data and aerial image.

References

[1] K. Khoshelham, C. Nardinocchi, C. Frontoni, A. Mancini, and P. Zingaretti, "Performance evaluation of automated approaches to building detection in multi-source aerial data," *ISPRS J. Photogramm. Remote Sens.*, vol. 65, no. 1, pp. 123–133, Jan. 2010.

[2] H. Miura and S. Midorikawa, "Updating GIS building inventory data using high-resolution satellite images for earthquake damage assessment: Application to Metro Manila, Philippines," *Earthquake Spectra*, vol. 22, no. 1, pp. 151–168, Feb. 2006.

[3] M. Awrangjeb, M. Ravanbakhsh, and C. Fraser, "Automatic detection of residential buildings using LIDAR data and multispectral imagery," *ISPRS J. Photogramm. Remote Sens.*, vol. 65, no. 5, pp. 457–467, Sep. 2010.

[4] H. Ruther, M. Hagai, and E. G. Mtalo, "Application of snakes and dynamic programming optimization in modeling of buildings in informal settlement areas," *ISPRS J. Photogramm. Remote Sens.*, vol. 56, no. 4, pp. 269–282, Jul. 2002.

[5] J. Peng, D. Zhang, and Y. Liu, "An improved snake model for building detection from urban aerial images," *Pattern Recognit. Lett.*, vol. 26, no. 5, pp. 587–595, Apr. 2005.

[6] M. Kabolizade, H. Ebadi, and S. Ahmadi, "An improved snake model for automatic extraction of buildings from urban aerial images and LiDAR data," *Comput. Environ. Urban Syst.*, vol. 34, no. 5, pp. 435–441, Aug. 2010.

[7] S. Ahmadi, M. J. Valadan Zoej, H. Ebadi, H. A. Moghaddam, and A. Mohammadzadeh, "Automatic urban building boundary extraction from high resolution aerial images using an innovative model of active contours," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 12, no. 3, pp. 150–157, Jun. 2010.

[8] N. Haala and C. Brenner, "Extraction of buildings and trees in urban environments," *ISPRS J. Photogramm. Remote Sens.*, vol. 54, no. 2, pp. 130–137, Jul. 1999.

[9] K. Zhang, J. Yan, and S.-C. Chen, "Automatic construction of building footprints from airborne LIDAR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 9, pp. 2523–2533, Sep. 2006.

[10] G. Miliaresis and K. Kokkas, "Segmentation and object-based classification for the extraction of the building class from LIDAR DEMs," *Comput. Geosci.*, vol. 33, no. 8, pp. 1076–1087, Aug. 2007.

[11] A. P. Nyaruhuma, M. Gerke, G. Vosselman, and E. G. Mtalo, "Verification of 2D building outlines using oblique airborne images," *ISPRS J. Photogramm. Remote Sens.*, vol. 71, pp. 62–75, Jul. 2012.

[12] E. Kwak and A. Habib, "Automatic representation and reconstruction of DBM from LiDAR data using recursive minimum bounding rectangle," *ISPRS J. Photogramm. Remote Sens.*, vol. 93, pp. 171–191, Jul. 2014.

[13] M. Gerke and J. Xiao, "Fusion of airborne laserscanning point clouds and images for supervised and unsupervised scene classification," *ISPRS J. Photogramm. Remote Sens.*, vol. 87, pp. 78–92, Jan. 2014.

[14] J. Secord and A. Zakhor, "Tree detection in urban regions using aerial LiDAR and image data," *IEEE Geosci. Remote Sens. Lett.*, vol. 4, no. 2, pp. 196–200, Apr. 2007.

[15] S. Cui, Q. Yan, and P. Reinartz, "Complex building description and extraction based on Hough transformation and cycle detection," *Remote Sens. Lett.*, vol. 3, no. 2, pp. 151–159, Mar. 2012.

[16] A. Sampath and J. Shan, "Building boundary tracing and regularization from airborne LiDAR point clouds," *Photogramm. Eng. Remote Sens.*, vol. 73, no. 7, pp. 805–812, Jul. 2007.

[17] G. Sohn and I. Dowman, "Data fusion of high-resolution satellite imagery and LiDAR data for automatic building extraction," *ISPRS J. Photogramm. Remote Sens.*, vol. 62, no. 1, pp. 43–63, May 2007.

[18] N. Ekhtari, M. J. Valadan Zoej, M. R. Sahebi, and A. Mohammadzadeh, "Automatic building extraction from LIDAR digital elevation models and worldview imagery," *J. Appl. Remote Sens.*, vol. 3, no. 1, p. 033571, Dec. 2009.

[19] M. Awrangjeb, C. Zhang, and C. S. Fraser, "Building detection in complex scenes thorough effective separation of buildings from trees," *Photogramm. Eng. Remote Sens.*, vol. 78, no. 7, pp. 729–745, Jul. 2012.

[20] T. T. Vu, F. Yamazaki, and M. Matsuoka, "Multi-scale solution for building extraction from LiDAR and image data," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 11, no. 4, pp. 281–289, Aug. 2009.

[21] X. Huang and L. Zhang, "Morphological building/shadow index for building extraction from high-resolution imagery over urban areas," *IEEE J. Sel. Topics. Appl. Earth Observ. Remote Sens.*, vol. 5, no. 1, pp. 161–172, Oct. 2011.

[22] Y. Tang, X. Huang, and L. Zhang, "Fault-tolerant building change detection from urban high-resolution remote sensing imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 10, no. 5, pp. 1060–1064, Jan. 2013.

[23] Y. Du, "Verification of tsunami reconstruction project by object-oriented building extraction from high resolution satellite imagery," M.S. thesis, Faculty of Geo-Information Science and Earth Observation, Int. Inst. GeoInf. Sci. Earth Obs., Enschede, The Netherlands, Mar. 2008.

[24] S. R. Lach, "Semi-automated DIRSIG scene modeling from 3D Lidar and passive imagery," Ph.D. dissertation, Chester F. Carlson Center for Imaging Science, Rochester Inst. Technol. (RIT), New York, NY, USA, May 2008.

[25] A. Turlapaty, B. Gokaraju, Q. Du, N. Younan, and J. Aanstoos, "A hybrid approach for building extraction from spaceborne multi-angular optical imagery," *IEEE J. Sel. Topics. Appl. Earth Observ. Remote Sens.*, vol. 5, no. 1, pp. 89–100, Jan. 2012.

[26] D. Mongus, N. Lukac, and B. Zalik, "Ground and building extraction from LiDAR data based on differential morphological profiles and locally fitted surfaces," *ISPRS J. Photogramm. Remote Sens.*, vol. 93, pp. 145–156, Jul. 2014.

[27] J. Shan and C. K. Toth, *Topographic Laser Ranging and Scaning*. Boca Raton, FL, USA: CRC Press/Taylor & Francis, ch. 13, pp. 390–420, 2008.

[28] J. Shan and A. Sampath, "Urban dem generation from raw LiDAR data: A labeling algorithm and its performance," *Photogramm. Eng. Remote Sens.*, vol. 71, no. 2, pp. 217–226, Feb. 2005.

[29] G. Sithole, "Segmentation and classification of airborne laser scanner data," Ph.D. dissertation, Faculty of Engineering and Geosciences, Delft Univ. Technol., Delft, Netherlands, May 2005.

[30] N. G. Silleos, T. K. Alexandridis, I. Z. Gitas, and K. Perakis, "Vegetation indices: Advances made in biomass estimation and vegetation monitoring in the last 30 years," *Geocarto Int.*, vol. 21, no. 4, pp. 21–28, Jan. 2006.

[31] J. W. Rouse, R. H. Haas, D. W. Deering, J. A. Schell, and J. C. Harlan, *Monitoring the Vernal Advancement and Retrogradation (Green Wave Effect) of Natural Vegetation*. College Station, TX, USA: Remote Sensing Center, Texas A and M Univ., 1974, ch. 3, pp. 23–57.

[32] G. S. Birth and G. McVey, "Measuring the color of growing turf with a reflectance spectroradiometer," *Agron. J.*, vol. 60, no. 6, pp. 640–643, 1968.

[33] D. Grigillo, M. K. Fras, and D. Petrovič, "Automatic extraction and building change detection from digital surface model and multispectral orthophoto," *Geod. Vestnik Assoc. Surv. Slovenia*, vol. 55, no. 1, pp. 28–45, Mar. 2011.

[34] A. R. Shahtahmassebi, Y. Ning, W. Ke, N. Moore, and S. Zhangquan, "Review of shadow detection and de-shadowing methods in remote sensing," *Chin. Geog. Sci.*, vol. 23, no. 4, pp. 403–420, Aug. 2013.

[35] K. L. Chung, Y. R. Lin, and Y. H. Huang, "Efficient shadow detection of color aerial images based on successive thresholding scheme," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 2, pp. 671–682, Feb. 2009.

[36] A. Mcandrew, *An Introduction to Digital Image Processing With Matlab*. Melbourne, VIC, Australia: School of Computer Science and Mathematics, Victoria Univ. Technology, ch. 11, pp. 191–213, 2004.

[37] C. Cortes and V. Vapnik, "Support vector networks," *Mach. learning*, vol. 20, no. 3, pp. 273–297, Sep. 1995.

[38] M. Rutzinger, F. Rottensteiner, and N. Pfeifer, "A comparison of evaluation techniques for building extraction from airborne laser scanning," *IEEE J. Sel. Topics. Appl. Earth Observ. Remote Sens.*, vol. 2, no. 1, pp. 11–20, Mar. 2009.

[39] L. Pereira and L. Janssen, "Suitability of laser data for DTM generation: A case study in the context of road planning and design," *ISPRS J. Photogramm. Remote Sens.*, vol. 54, no. 4, pp. 244–253, Sep. 1999.

[40] S. E. Reutebuch, R. J. McGaughey, H. E. Andersen, and W. W. Carson, "Accuracy of a high-resolution LiDAR terrain model under a conifer forest canopy," *Can. J. Remote Sens.*, vol. 29, no. 5, pp. 527–535, Oct. 2003.

[41] C. J. C. Burges, "A tutorial on support vector machines for pattern recognition," *Data Min. Knowl. Discovery*, vol. 2, no. 2, pp. 121–167, Jun. 1988.

[42] F. Rottensteiner, G. Sohn, M. Gerke, J. D. Wegner, U. Breitkopf, and J. Jung, "Results of the ISPRS benchmark on urban object detection and 3D building reconstruction," *ISPRS J. Photogramm. Remote Sens.*, vol. 93, pp. 256–271, Jul. 2014.

[43] M. Cramer, "The DGPF test on digital aerial camera evaluation—-Overview and test design," *Photogramm. Fernerkundung Geoinf.*, vol. 2, pp. 73–82, Jan. 2010.

Asghar Zarea received the B.Sc. degree from the University of Tabriz, Tabriz, Iran, and the M.Sc. degree in geomatic engineering and remote sensing from K.N. Toosi University of Technology, Tehran, Iran, in 2011 and 2013, respectively.

He is currently teaching at Azad University of Mamaghan, Mamaghan, Iran, where he is the responsible for the university courses entitled Digital Terrain Model, Image Processing, Introduction to Remote Sensing, and Photogrammetry. He is the author of two journal papers in national space journals and the author or coauthor of five conference papers. His research interests include airborne light detection and ranging mapping, image processing, and feature detection.

Ali Mohammadzadeh received the B.Sc. and M.Sc. degrees in geodesy and geomatics engineering, and the Ph.D. degree in remote sensing from K.N.Toosi University of Technology, Tehran, Iran, in 2000, 2002, and 2009, respectively.

Since January 2010, he is an Assistant Professor and the Head of the Department of Photogrammetry and Remote Sensing, K.N.Toosi University of Technology. Also, he is a responsible for the university courses entitled Management of Remote Sensing Data, Concepts of Laser Scanners, and Artificial Intelligence at the respective University. He has over 30 published conference papers and 12 published journal papers. His research interests include dust detection, remote sensing image analysis, artificial intelligence, LiDAR image analysis, and feature extraction.