Contribution of Normalized DSM to Automatic Building Extraction from HR Mono Optical Satellite Imagery

Umut Gunes Sefercik1*, Serkan Karakis1, Caglar Bayik1, Mehmet Alkan2 and Naci Yastikli2

1Bulent Ecevit University, Department of Geomatics Engineering, Zonguldak, Turkey
2Yildiz Technical University, Department of Geomatics Engineering, Istanbul, Turkey

*Corresponding author, e-mail address: ugsefercik@hotmail.com

Abstract
Building extraction from high resolution (HR) satellite imagery is one of the most significant issue for remote sensing community. Manual extraction process is onerous and time consuming that’s why the improvement of the best automation is a crucial topic for the researchers. In this study, we aimed to expose the significant contribution of normalized digital surface model (nDSM) to the automatic building extraction from mono HR satellite imagery performing two-step application in an appropriate study area which includes various terrain formations. In first step, the buildings were manually and object-based automatically extracted from ortho-rectified pan-sharpened IKONOS and Quickbird HR imagery that have 1 m and 0.6 m ground sampling distances (GSD), respectively. Next, the nDSM was created using available aerial photos to represent the height of individual non-terrain objects and used as an additional channel for segmentation. All of the results were compared with the reference data, produced from aerial photos that have 5 cm GSD. With the contribution of nDSM, the number of extracted buildings was increased and more importantly, the number of falsely extracted buildings occurred by automatic extraction errors was sharply decreased, both are the main components of precision, completeness and overall quality.

Keywords: automatic building extraction, high resolution satellite imagery, normalized digital surface model (nDSM), precision, completeness, quality.

Introduction
In recent decades, optical space-borne remote sensing technologies have been improved rapidly and at the end of 1990s, with the innovation of IKONOS, commercial optical satellites offering high resolution (HR) (≤1 m) have become utilizable. By the advantage of HR, optical satellite data were started to use for building extraction more than before [Theng, 2006; Sportouche et al., 2009; Ehrlich et al., 2012; Benarchid et al., 2013; Bhadauria et al., 2013]. Overall, two basic methods are employed for building extraction as manual and
automatic. The automatic extraction process is definitely faster and time saving in comparison with manual extraction process and additionally enables the extraction of vector semantic data [Baltsavias et al., 2001]. Furthermore, object-based automatic extraction data can be easily transferred into computer aided design (CAD) and geographic information system (GIS) based software and applications. However, automatic building extraction includes unfavorable situations caused by imaging radiometry of optical sensors depending on the similar spectral reflectance of the ground objects [Shaw and Burke, 2003; Pedelty et al., 2004; Bedard et al., 2012]. The roof types and colors are various for settlements depending on the different construction techniques and used materials. The similar spectral reflectance of other ground objects with different characterized roofs causes missing buildings and especially falsely extracted buildings on the ground as results of automatic extraction. As known, the space-borne imagery are usually used for urban applications such as monitoring of urban development and detection of illegal construction. Especially in large urban areas, these applications are realized by temporal change detection, calculated in the basis of automatic building extraction [Baiocchi et al., 2011]. At this point, the falsely extracted buildings, occurred by the segmentation errors, complicate the detection of changes in urban areas and lead to erroneous results. One of the main causes of this situation is two dimensional (2D) segmentation in automatic extraction process.

In this study, we aimed to expose the significant contribution of normalized digital surface model (nDSM) on the solution of problems caused by the lack of height information. A digital surface model (DSM) explains the three dimensional (3D) earth surface including all of terrain and non-terrain objects such as buildings, vegetation, forest, and roads. On the other hand, a digital terrain model (DTM) represents only the 3D bare earth topography. Based out of here, nDSM is a differential model, calculated by the subtraction of DTM and DSM of a common area. This concept is mainly used in forestry discipline for detecting the height of trees and forest stands [Stereńczak et al., 2008; Smreček, 2012; Sefercik and Ateşoğlu, 2013]. The potential of nDSM with respect to building extraction was explored only by a few authors using airborne laser scanning (ALS) data [Ma, 2005; Yu et al., 2010; Ahearn and Ahn, 2011]. However, ALS is a young technique and not available for a large portion of the earth yet. As distinct from ALS-based studies, we exhibited the significant contribution of nDSM, produced from aerial imagery to the object-based automatic building extraction from space-borne imagery regarding standard quantitative measures as precision (correctness), completeness, and overall quality [Heipke et al., 1997; Karantzalos and Paragios, 2009]. The case study was realized using ortho-rectified pan-sharpened IKONOS and Quickbird satellite imagery that have 1 m and 0.6 m ground sampling distances (GSD) respectively in the City of Zonguldak, Turkey having various land classes and rough terrain structure which are very beneficial to test the proposed technique.

Considering the purposes, the paper is organized as follows: the next section describes the study area and used materials. This is followed by the methodologies used for the processing of satellite images containing geometric correction, pan-sharpening, and ortho-rectification and for the building extraction. In the next section, the results and discussions are presented followed by the conclusion of the investigation.

**Study area and datasets**
As known, one of the most significant problems for space-borne imagery is the inclined topographies that’s why fully flat areas are always misleading to exhibit the performance
of proposed techniques. To demonstrate the contribution of our technique, we consciously preferred a study area that includes rough terrain and various land classes. The study area covers 1 km × 1 km in the City of Zonguldak, located on the North-West side of Turkey. It is adjacent with Black Sea shoreline in two borders mostly having vertical cliffs with more than 90° inclination angle. The average terrain inclination is 12% and the elevation of bare topography varies from sea level up to 147 m. The large part of area (~85%) is protected and new building construction has been prohibited for more than fifty years. Five different land classes are available as open, built-up (apartments, education and training institutions etc.), forest, roads, and rocks along the shoreline. Figure 1 shows the aerial photo of the study area (Fig. 1a), land classes (Fig. 1b), color-coded DTM (Fig. 1c), and the distribution of terrain slope (Fig. 1d). The reference aerial photos of the study area were acquired by large scale photogrammetry in 2009. They have 5 cm GSD and ±15 cm horizontal geo-location accuracy. The bundle panchromatic (PAN) + multispectral (MS) mono IKONOS and Quickbird satellite imagery, used for the implementation were acquired in 2008 and 2004, respectively. The time interval between the acquisition dates was neglected because of building construction prohibition in the study area. Figure 2 illustrates the pan-sharpened IKONOS and Quickbird images and following Table 1 gives information about the characteristics of them.

![Figure 1 - Study area (a), land classes (b), color-coded DTM (c) and the distribution of terrain slope (d).]
Table 1 - Characteristics of IKONOS and Quickbird imagery.

| Property          | IKONOS                  | Quickbird              |
|-------------------|-------------------------|------------------------|
| Acquisition year  | 2008                    | 2004                   |
| Origin            | USA                     | USA                    |
| Resolution        | 1 m PAN, 4 m MS         | 0.61 m PAN, 2.44 m MS  |
| Altitude          | 681 km                  | 450 km                 |
| Inclination angle | 98º.1                   | 97º.2                  |
| Swath width       | 11.3 km                 | 16.5 km                |
| Digital recording format | 11 bit                  | 11 bit                |

Methodologies for image processing and building extraction

The followed methodology for image processing and building extraction is summarized in a six-step work flow diagram as shown in Figure 3.
As known, the raw space-borne images are distributed with an approximate orientation. The absolute position accuracy of images based on the orbital information and position and orientation systems of satellites, are not satisfactory for applications that need high precision [Heipke, 1997; Büyüksalih et al., 2004; Toutin, 2004]. The geometric position accuracy of images should be improved according to the requirements of applications [Buiten and Van Putten, 1997; Jacobsen, 2004; Jacobsen et al., 2005]. Starting from this point, each image was oriented using rigorous satellite orbital modelling (SOM) [Toutin, 2004] utilizing uniformly distributed 3 independent check points (CPs) and 11 ground control points (GCPs), collected by real time kinematic GPS observations on the ground having minimum 3 cm position accuracy both for horizontal and vertical directions. The independent CPs for each image were located as one in the midpoint of the image, two at the inside and outside of predetermined study area. While GCPs were used for the initial geo-referencing, CPs were used for the assessment check. Table 2 shows the root mean square error (RMSE) of used CPs and GCPs.

### Table 2 - RMSE of used CPs and GCPs for geometric correction.

| Image      | GCPs       | CPs       | Total     |
|------------|------------|-----------|-----------|
|            | RMSE X (pix) | RMSE Y (pix) | RMSE (pix) | RMSE X (pix) | RMSE Y (pix) | RMSE (pix) |
| IKONOS PAN | 11         | 0.18      | 0.26      | 0.31       | 3           | 0.54       | 0.98       | 1.12       |
| IKONOS MS  | 11         | 0.33      | 0.47      | 0.57       | 3           | 0.80       | 0.74       | 1.09       |
| Quickbird PAN | 11     | 0.39      | 0.21      | 0.45       | 3           | 0.37       | 0.33       | 0.50       |
| Quickbird MS | 11      | 0.38      | 0.30      | 0.48       | 3           | 0.23       | 0.40       | 0.46       |

Next, HR fused color images were generated by the combination of corrected PAN and MS images using pan-sharpening technique utilizing UNB (University of New Brunswick) algorithm [Zhang, 2002; Padwick et al., 2010]. Pan-sharpening method makes it possible to benefit from the sensors’ spectral capabilities simultaneously with its high spatial resolution. In the literature, many algorithms such as principal component analysis (PCA), hue intensity saturation (HIS), Wavelet and Gram Schmidt are available for pan-sharpening [Nikolakopoulos, 2008]. The PAN images are achieved in the 0.4 - 0.9 µm wavelength interval of electromagnetic spectrum and very sensitive for green color due to including infrared band. The most significant advantage of UNB in comparison with other algorithms is the presentation of better results for green band of images. It works in statistical basis and uses the least squares method to calculate the contribution of each imaging band to the final product to find the best coherence between the combined imaging bands’ grey values and reduce the spectral distortion independent to the data set.

To make an appropriate object recognition, extraction or classification using satellite data, the ortho-images must be used. Accordingly, the ortho-images were generated from pan-sharpened images using 1 m original gridded DTM derived from reference aerial photos [Hofmann, 2001; Mena, 2003; Alkan et al., 2006; Topan et al., 2009]. For geometric correction, pan-sharpening and ortho-image generation, PCI Geomatica v9.1.1 software was used. The buildings were extracted from ortho-rectified images by manual and
automatic methods using NetCAD v5.2 and eCognition v4.0.6 software respectively. For the automatic extraction of buildings, an object-based approach was selected based on the combination of spectral (color) and shape heterogeneity changes [Baatz et al., 2004]. In object-based classification, the heterogeneity is calculated based on fused adjacent objects in a stable direction (top-down, bottom to up etc.) limiting by the pre-determined scale parameter. The fused objects have to continue as the first object (obj\textsubscript{1}) and the second object (obj\textsubscript{2}). For instance, the fusion of obj\textsubscript{1} and obj\textsubscript{2} will be named as obj\textsubscript{1} for the next merge and merge with a further obj\textsubscript{2} again. The spectral heterogeneity of an object is calculated by the standard deviation of concerning pixel numbers (colors). The shape heterogeneity consists of two components as compactness and smoothness. The compactness parameter represents the closeness of the pixels clustered in an object by comparing it to a circle while the smoothness parameter describes the similarity between the image object borders and a perfect square [UTSA, 2013]. These parameters and overall heterogeneity can be calculated by following equations;

\[
\Delta h_{sp} = \sum_{i=1}^{N} W_i \left( n_{\text{merge}} \sigma_i^{\text{merge}} - \left( n_{\text{obj1}} \sigma_i^{\text{obj1}} + n_{\text{obj2}} \sigma_i^{\text{obj2}} \right) \right) [1]
\]

\[
\Delta h_{cm} = n_{\text{merge}} \frac{l_{\text{merge}}}{\sqrt{n_{\text{merge}}}} - \left( \frac{n_{\text{obj1}} l_{\text{obj1}}}{\sqrt{n_{\text{obj1}}}} + \frac{n_{\text{obj2}} l_{\text{obj2}}}{\sqrt{n_{\text{obj2}}}} \right) [2]
\]

\[
\Delta h_{sm} = n_{\text{merge}} \frac{l_{\text{merge}}}{b_{\text{merge}}} - \left( \frac{n_{\text{obj1}} l_{\text{obj1}}}{b_{\text{obj1}}} + \frac{n_{\text{obj2}} l_{\text{obj2}}}{b_{\text{obj2}}} \right) [3]
\]

\[
\Delta h_{sh} = W_{cm} \Delta h_{cm} + (1 - W_{cm}) \Delta h_{sm} [4]
\]

\[
\Delta h_{\text{overall}} = (1 - W_{sh}) \Delta h_{sp} + W_{sh} \Delta h_{sh} [5]
\]

where \( \Delta h_{\text{overall}}, \Delta h_{sp}, \Delta h_{sh}, \Delta h_{cm}, \Delta h_{sm} \) are the overall, spectral, shape, compactness, and smoothness heterogeneity respectively. \( N \) is the number of channels of the segmented satellite image, \( W \) represents the channel weight, \( n \) describes the number of pixels belonging to the object 1 and 2 (obj\textsubscript{1}, obj\textsubscript{2}), \( \sigma \) is the standard deviation, \( l \) and \( b \) are the factual border length and the perimeter of the bounding obj\textsubscript{1}, obj\textsubscript{2} [Tian and Chen, 2007].

The classification was performed based on object properties considering four different land classes as buildings, roads, sea, and the vegetation. To separate the external land classes except buildings and roads, normalized difference vegetation index (NDVI) and near infrared and green band differences were utilized [Sohn and Dowman, 2007]. Table 3 shows the segmentation parameters in four levels used for most successful classification.
The optimal values for segmentation parameters are defined by the operator considering used image characteristics (resolution, viewing geometry, radiometry, distortions, etc.) and object properties (length, width, roof types, etc.) in the study area and also resulting segmentation outcome. The meanings of the segmentation parameters are summarized as follows [Hofmann, 2001; Benz et al., 2003]:

- **level**: This parameter determines whether a new generated image level will either overwrite a current level or whether the generated objects shall become sub- or super-objects of a still existing level. The order of generating the levels affects the objects’ shape (top-down, bottom-up segmentation);
- **scale**: As mentioned before, heterogeneity is calculated based on the fusion of adjacent objects. The scale parameter describes a threshold value to stop the fusion of objects and segmentation. That means, if larger scale is used more objects are fused in segmentation and larger objects occur;
- **color and shape**: The influence of spectral and shape homogeneity is adjusted by these parameters in segmentation. The total of these parameters is equal to 1. Accordingly, high spectral homogeneity value means less shape homogeneity influence in segmentation;
- **compactness and smoothness**: Using these parameters, the user defines whether the objects shall become more compact (fringed) or more smooth;
- **weight of image channels**: This parameter determines the weight of an image channel on the segmentation. For the images that have comparable channels in size and content such as IKONOS and Quickbird each channel should be weighted equally.

### Table 3 - Segmentation parameters and applied thresholds for IKONOS and Quickbird imagery.

| Segmentation levels | Scale | Color | Shape | Compactness | Smoothness | Weights       |
|---------------------|-------|-------|-------|-------------|------------|---------------|
| **IKONOS**          |       |       |       |             |            |               |
| 1                   | 20    | 0.3   | 0.7   | 0.5         | 0.5        | 1,1,1,1,1     |
| 2                   | 40    | 0.3   | 0.7   | 0.5         | 0.5        | 1,1,1,1,1     |
| 3                   | 60    | 0.3   | 0.7   | 0.5         | 0.5        | 1,1,1,1,1     |
| 4                   | 20    |       |       |             |            | 1,1,1,1,0,1   |
| **Quickbird**       |       |       |       |             |            |               |
| 1                   | 20    | 0.1   | 0.9   | 0.5         | 0.5        | 1,1,1,1,1     |
| 2                   | 40    | 0.1   | 0.9   | 0.5         | 0.5        | 1,1,1,1,1     |
| 3                   | 60    | 0.1   | 0.9   | 0.5         | 0.5        | 1,1,1,1,1     |
| 4                   | 20    |       |       |             |            | 1,1,1,1,0,1   |

As can be seen in Table 3, different values were used for color and shape heterogeneity for IKONOS and Quickbird images because of geometric (resolution, viewing angle etc.) and radiometric differences (Fig. 4). In addition to lower resolution (1 m), the radiometry of raw IKONOS image is not as good as Quickbird and includes single white spots that are marked with circles in Figure 4. The large portion of the white spots was eliminated using median filtering. However, the residuals of them effect the shape heterogeneity and cause problems on the combination of the details in segmentation. That’s why we increased the color heterogeneity for IKONOS and eliminate the influence of these single white spots which cannot significantly change the average of grey values when combining with neighbour pixels.
In 2D automatic extraction process, the spectral reflectance and the shape of the ground objects are considered for segmentation. However, especially in underdeveloped and developing countries, regular settlement plans and architecture are not available in most cases. Associated with construction methods and economic reasons, different kinds of roofs (tile, concrete, wooden, steel, zinc, sheet metal etc.) are used for the buildings. These different roof types always create difficulties for segmentation algorithms because of dissimilar spectral reflectance and shapes. In addition, other land classes may have similar spectral reflectance and shape with the roofs. The best sample for this situation is the roads and concrete roofs that are often nested in segmentation. To avoid these challenges by providing 3D information, a 1 m gridded nDSM was generated for the study area with LISA 5.1 software. This is a differential height model of the DSM and DTM [6] serving up the pure non-terrain object heights. The DSM and the DTM were generated from aerial photos by photogrammetric assessment based on optical stereoscopy using Z/I imaging and Microstation software [Pulighe and Fava, 2013]. During photogrammetric assessment, the points were collected on the stereo-model under two layers as terrain and non-terrain. The DSM was generated using both terrain and non-terrain points. For DTM generation, only the non-terrain points that represent the bare topography were employed. The absolute vertical accuracies of both models are in between 10 cm and 1 m. To provide the most realistic topographic description and aimed grid spacing (1 m) for final DSM and DTM, a dense number of points were collected (approx. 14000). Thus, the negative influence of interpolation in the generation of DSM and DTM was minimized. Considering the changing topographic structure in the study area, a common interpolation method ‘triangulation’ was preferred when generating the 3D models. Figure 5 shows the generated DSM (Fig. 5a), DTM (Fig. 5b), and built-up areas on the 3D nDSM (Fig. 5c) as a wire frame where the maximum building height is 42.6 m. The exaggeration factor was selected as ‘2’ for better interpretation.

\[ nDSM = DSM - DTM \] [6]
As seen in Figure 5, the buildings cannot be described with sharp lines by nDSM. The main cause for this situation is the original grid spacing (1 m) of the DSM and DTM that were used for nDSM generation and related interpolation. As known, the main parameter for most realistic DSM and DTM generation using optical stereoscopy is the number of proper matching and height points that are marked by the operator. The heights of the entire included points in requested grid interval of a DSM or a DTM are automatically calculated by interpolation which decreases the vertical accuracies of them [Passini and Jacobsen, 2007; Sefercik and Alkan, 2009]. The most significant factor that minimizes the negative influence of interpolation is measured point density for the generation area. This means, considering the advantage of point density, the vertical accuracy and building representation quality of a generated nDSM from ALS data will be better than generated from aerial photos. However, ALS is a young technique and not available for many countries where aerial photos exist.
At the last stage of application, automatic building extraction was performed with IKONOS and Quickbird pan-sharpened images using the contribution of nDSM as an additional channel for the segmentation with the weight of ‘0.2’ for all segmentation levels. The results were compared with the first extraction regarding two main components as the number of extracted and falsely extracted buildings.

**Results and discussion**

The manual building extraction results of HR IKONOS and Quickbird pan-sharpened imagery are shown in Figure 6 and following Table 4. The extracted buildings (red) were overlapped with the buildings derived from reference aerial photos (blue) for better interpretation. For the clear presentation, all extraction results, presented in this section were grouped using the common classification terms true positives (TP), false positives (FP), and false negatives (FN) [Clode et al., 2004; Mancini et al., 2009]. As known, in classification applications, TP, FP, and FN are achieved when a test correctly reports a positive result, when a test falsely reports a positive result, and when a test incorrectly gives a negative result, respectively. Accordingly, truly extracted buildings are TP, falsely extracted buildings are FP, and unidentified buildings are FN in this study. In the Table 4, TP and FN are shown for manual building extraction. In manual extraction, no FP result are achieved that’s why this result does not exist in Table 4.

![Figure 6 - Manually extracted buildings from HR IKONOS (a) and Quickbird imagery (b).](image)

**Table 4 - Manual building extraction results.**

| Data     | TP  | FN  | Extraction (%) |
|----------|-----|-----|----------------|
| Reference| 362 | 0   | 100            |
| IKONOS   | 338 | 24  | 93.3           |
| Quickbird| 351 | 11  | 97.0           |
As seen in Table 4, 362 buildings are available in the study area and most of them can be manually extracted from both satellite images. Although, the results are very similar, the Quickbird is one step ahead in the detection of small buildings by the advantage of higher resolution and better radiometry.

Considering the availability of buildings in IKONOS and Quickbird images, demonstrated by Figure 6 and Table 4, the majority of problems in automatic extraction performance will arise because of segmentation. Figure 7 and following Table 5 show the object-based automatic building extraction results. In Table 5, in addition to TP and FN, the FP are represented. The precision (correctness), completeness, and the overall quality were calculated as:

\[
P = \frac{TP}{TP + FP} \quad [7]
\]

\[
C = \frac{TP}{TP + FN} \quad [8]
\]

\[
Q = \frac{TP}{TP + FP + FN} \quad [9]
\]

where \( P, C, Q \) are the precision, completeness, and quality of automatic building extraction, respectively.

![Figure 7 - Automatic extracted objects from IKONOS (a) and Quickbird imagery (b).](image)
Table 5 - Object-based automatic building extraction results without nDSM.

| Data        | TP | FN | FP | Precision (%) | Completeness (%) | Quality (%) |
|-------------|----|----|----|---------------|------------------|-------------|
| Reference   | 362| 0  | 0  | 100           | 100              | 100         |
| IKONOS      | 299| 63 | 126| 70.1          | 82.6             | 61.3        |
| Quickbird   | 329| 33 | 163| 66.9          | 90.9             | 62.7        |

As can be seen in Table 5, falsely extracted building numbers in IKONOS and Quickbird are 126 and 163, respectively that have a considerable influence on precision and overall quality of automatic object extraction. As previously mentioned, the main causes of falsely extracted buildings are similar spectral reflectance and shape of other non-terrain objects. For instance, the road of breakwater in city port and boats were extracted as a building in the images (white rectangles in Fig. 7). The numbers of falsely extracted buildings were manually calculated considering available buildings on the reference data. This calculation was made just to show the contribution of nDSM on the next step because a manual calculation of falsely extracted buildings is meaningless and impossible in large application areas. As mentioned earlier, the main cause of falsely extracted buildings is the lack of height information in 2D automatic extraction process. At this point, the significant contribution of correct height information provided from nDSM was demonstrated by repeated whole automatic process. Figure 8 and Table 6 present the results of automatically extracted buildings using the nDSM. In Figure 8, it can be easily distinguished that there is a clear improvement on the results in comparison with automatic extraction without nDSM. The individual buildings are clearer and a sharp reduction is monitored on the falsely extracted building areas. For an instance, white rectangles in Figure 7 and 8 can be compared.
Regarding Table 6, the contribution of nDSM is mostly effect to the precision and overall quality of the automatic building extraction from both space-borne HR data. By the advantage of better resolution and radiometry, the improvements of Quickbird results are narrowly more than IKONOS. The falsely extracted building numbers were decreased from 126 to 32 and 163 to 34 where the extracted buildings were increased from 299 to 300 and 329 to 338 in order of IKONOS and Quickbird. Accordingly, the precision of the extraction increased from 70.1% to 90.4% and 66.9% to 90.9% and the overall quality increased from 61.3% to 76.1% and 62.7% to 85.4% for IKONOS and Quickbird HR imagery, respectively. The completeness was also improved by nDSM but not in the similar ratios with precision and overall quality. It increased from 82.6% to 82.9% and 90.9% to 93.4% in order of IKONOS and Quickbird.

In addition to the aforementioned calculations, areal accuracy of extracted buildings were investigated based on the calculation of areal differences between extracted buildings from tested images and reference data. This is one of the common methods to introduce the accuracy of extracted buildings [Shiravi et al., 2012; GRC, 2014]. Table 7 presents the total areas of extracted buildings from space-borne imagery with and without nDSM and the areal differences from the reference data. With the contribution of nDSM, the areal difference between IKONOS and reference data decreased as 48%. This result demonstrates that the applicants should avoid using 2D methods for automatic building extraction from mono IKONOS images in rough topographies. For Quickbird imagery, the areal difference from reference was also decreased as 13%.

Overall, the results verify that the generated nDSM from available aerial photos is very favorable especially for eliminating the large portion of falsely extracted buildings (numerical and areal) and improves the precision, completeness, and the overall quality of object-based automatic building extraction from mono space-borne HR imagery.

### Table 6 - Object-based automatic building extraction results using nDSM.

| Data       | TP  | FN  | FP  | Precision (%) | Completeness (%) | Quality (%) |
|------------|-----|-----|-----|---------------|------------------|-------------|
| Reference  | 362 | 0   | 0   | 100           | 100              | 100         |
| IKONOS     | 300 | 62  | 32  | 90.4          | 82.9             | 76.1        |
| Quickbird  | 338 | 24  | 34  | 90.9          | 93.4             | 85.4        |

### Table 7 - Extracted building areas with and without nDSM.

| Data           | Extracted Building Area (m²) | Areal Difference from Reference (m²) | Areal Difference from Reference (%) |
|----------------|-----------------------------|-------------------------------------|------------------------------------|
| Reference      | 102851                      | 0                                   | 0.00                               |
| IKONOS without nDSM | 163303                | 60452                               | 58.78                              |
| Quickbird without nDSM | 125107                  | 22256                               | 21.64                              |
| IKONOS with nDSM     | 113388                    | 10537                               | 10.25                              |
| Quickbird with nDSM  | 111632                    | 8781                                | 8.54                               |
Conclusion
In this study, an alternative technique was proposed for the improvement of precision, completeness, and overall quality of automatic building extraction from space-borne HR mono satellite images. The mono images were selected because of lower cost and easy processing in comparison with stereos. An nDSM that provides correct height information of non-terrain objects was generated from available aerial photos and used as an additional channel for segmentation. In the literature, all of nDSM based studies use ALS data which is not available for underdeveloped or developing Countries. In contrast to ALS data, the aerial photos are available for almost entire Countries which produce their own topographic maps. The geometric and radiometric properties of space-borne imagery have vital importance on the selection of spectral and shape heterogeneity thresholds in object-based automatic building extraction. For IKONOS imagery, due to lower geometric and radiometric conditions in comparison with Quickbird, a higher color threshold was preferred. In the object-based automatic building extraction, in addition to extracted building number, falsely extracted building number is very crucial. The falsely extracted buildings, occurred by the segmentation errors, complicate the detection of changes in urban areas and lead to erroneous results. In the study, with the significant contribution of nDSM, a considerable improvement was provided on the object-based automatic building extraction especially eliminating falsely extracted buildings (numerical and areal). Through, the precision and overall quality of application was increased more than 40% for both IKONOS and Quickbird images. The completeness was also improved with smaller percentages against precision and overall quality.

Acknowledgements
Thanks are going to Zonguldak Municipality to provide IKONOS 2008 high resolution satellite images and Dr. Wilfred Linder and Dr. Karsten Jacobsen to supply LISA software.

References
Ahearn S.C., Ahn H.J. (2011) - Quality assurance and potential applications of a high density lidar data set for the city of New York. ASPRS Annual Conference, Newyork, USA, pp. 9.
Alkan M., Marangoz A.M., Karakis S., Büyüksalih G. (2006) - Verification of automatic and manual road extraction methods using quickbird imagery. ISPRS Archives, 36 (W41), pp. 4.
Baatz M., Benz U., Dehghani S., Heynen M., Höltje A., Hofmann P., Lingenfelder I., Mimler M., Sohlbach M., Weber M., Willhauck G. (2004) - Ecognition professional: user guide 4. Definiens Imaging, Munih, Germany, pp.486.
Baiocchi V., Brigante R., Radicioni F. (2010) - Three-dimensional multispectral classification and its application to early seismic damage assessment. European Journal of Remote Sensing, 42: 49-65. doi: http://dx.doi.org/10.5721/ItJRS20104234.
Baltsavias E.P., Gruen A., Van Gool L. (2001) - Automatic extraction of man-made objects from aerial and space images (III). Taylor & Francis. A. A. Balkema Publishers, pp. 425. ISBN: 90-5809-252-6.
Bedard N., Hagen N., Gao L., Tkaczyk T.S. (2012) - *Image mapping spectrometry: calibration and characterization*. Optical Engineering, 51 (11): pp.14. doi: http://dx.doi.org/10.1117/1.OE.51.11.111711.

Benarchid O., Raissouni N., El Adib S., Abbous A., Azyat A., Achhab N.B., Lahraoua M., Chahboun A. (2013) - *Building extraction using object-based classification and shadow information in very high resolution multispectral images, a case study: Tetuan, Morocco*. Canadian Journal on Image Processing and Computer Vision, 4 (1): 1-8.

Benz U.C., Hofmann P., Willhauck G., Lingenfelder I., Heynen M. (2004) - *Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information*. ISPRS Journal of Photogrammetry and Remote Sensing, 58 (3-4): 239-258. doi: http://dx.doi.org/10.1016/j.isprsjprs.2003.10.002.

Bhadauria A., Bhadauria H., Kumar A. (2013) - *Building extraction from satellite images*. IOSR Journal of Computer Engineering, 12 (2): 76-81. doi: http://dx.doi.org/10.9790/0661-1227681.

Buiten H., Van Putten B. (1997) - *Quality assessment of remote sensing image registration-analysis and testing of control point residuals*. ISPRS Journal of Photogrammetry and Remote Sensing, 52: 57-73. doi: http://dx.doi.org/10.1016/S0924-2716(97)83001-8.

Büyüksalih G., Oruc M., Jacobsen K. (2004) - *Precise georeferencing of rectified high resolution space images*. XXth ISPRS Congress, Istanbul, Turkey, 35: 184-188.

Clode S., Kootsookos P., Rottensteiner F. (2004) - *The automatic extraction of roads from LiDAR data*. Proc. XXth ISPRS Congress, Istanbul, Turkey, 35: 231-236.

Ehrlich D., Kemper T., Blaes X., Soille P. (2012) - *Extracting building stock information from optical satellite imagery for mapping earthquake exposure and its vulnerability*. Natural Hazards, 68: 79-95. doi: http://dx.doi.org/10.1007/s11069-012-0482-0.

GRC (2014) - *Automated building extraction and reconstruction from LIDAR data*. http://www.grc.missouri.edu/icrestprojarchive/NASA/FeatureExtraction-buildings/Building%20Extraction.pdf, January 2014, pp.27.

Heipke C. (1997) - *Automation of interior, relative, and absolute orientation*. ISPRS Journal of Photogrammetry and Remote Sensing, 52: 1-19. doi: http://dx.doi.org/10.1016/S0924-2716(96)00029-9.

Heipke C., Mayer H., Wiedemann C., Jamet O. (1997) - *Evaluation of automatic road extraction*. ISPRS Archives, (32) 3-2W3, 47-56.

Hofmann P. (2001) - *Detecting urban features from IKONOS data using an object-oriented approach*. In: Proceedings of the 1st Annual Conference of the Remote Sensing & Photogrammetry Society, pp. 28-33.

Hofmann P. (2001) - *Detecting informal settlements from Ikonos image data using methods of object oriented image analysis-an example from Cape Town (South Africa)*. C. Jürgens (Ed.), Remote Sensing of Urban Areas / Fernerkundung in urbanen Räumen, Regensburger Geographische Schriften, pp. 107-118.

Jacobsen K. (2004) - *Geometry of satellite images–calibration and mathematical models*. Korean Society of Remote Sensing (Hrsg.): ISPRS international conference, Jeju, pp. 182-185.

Jacobsen K., Büyüksalih G., Topan H. (2005) - *Geometric models for the orientation of high resolution optical satellite sensors*. IAPRS, 36: 1.

Karantzalos K., Paragios N. (2009) - *Competing 3D priors for object extraction in remote
sensing data, In: Stilla U., Rottensteiner F., Paparoditis N. (Eds) CMRT09. IAPRS, Paris, France, 38 (3/W4): 127-132.

Ma R. (2005) - DEM generation and building detection from Lidar data. Photogrammetric Engineering and Remote Sensing, 71: 847-854. doi: http://dx.doi.org/10.14358/PERS.71.7.847.

Mancini A., Frontoni E., Zingaretti P. (2009) - A winner takes all mechanism for automatic object extraction from multi-source data. Geoinformatics, 17th International Conference, pp. 1-6.

Mena J.B. (2003) - State of the art on automatic road extraction for GIS update: a novel classification. Pattern Recognition Letters, 24: 3037-3058. doi: http://dx.doi.org/10.1016/S0167-8655(03)00164-8.

Nikolakopoulos K.G. (2008) - Comparison of nine fusion techniques for very high resolution data. Photogrammetric Engineering and Remote Sensing, 74: 647. doi: http://dx.doi.org/10.14358/PERS.74.5.647.

Padwick C., Deskevich M., Pacifici F., Smallwood S. (2010) - WorldView-2 pan-sharpening. ASPRS 2010 Annual Conference, San Diego, California, pp.14.

Passini R., Jacobsen K. (2007) - Accuracy analysis of SRTM height models. ASPRS Annual Conference, Tampa, pp. 25-29.

Pedelty J.A., Morisette J.T., Smith J.A. (2004) - Comparison of Landsat-7 enhanced thematic mapper plus (ETM+) and earth observing one (EO-1) advanced land imager. Optical Engineering, 43: 954-962. doi: http://dx.doi.org/10.1117/1.1651556.

Pulighe G., Francesco F. (2013) - DEM extraction from archive aerial photos: accuracy assessment in areas of complex topography. European Journal of Remote Sensing, 46: 363-378. doi: http://dx.doi.org/10.5721/EuJRS20134621.

Sefercik U., Atesoglu A. (2013) - Uydu verileri kullanılarak meşcere boyu belirlenmesine yönelik yeni bir yaklaşım. TUFUAB VIIth Technical Symposium, Trabzon, Turkey, pp. 5.

Sefercik U.G., Alkan M. (2009) - Advanced analysis of differences between C and X bands using SRTM data for mountainous topography. Journal of the Indian Society of Remote Sensing, 37: 335-349. doi: http://dx.doi.org/10.1007/s12524-009-0044-4.

Shaw G.A., Burke H.K. (2003) - Spectral imaging for remote sensing. Lincoln Laboratory Journal, 14: 3-28.

Shiravi S., Zhong M., Beykaei S.A. (2012) - Accuracy assessment of building extraction using LIDAR data for urban planning transportation applications, New Brunswick, Canada, pp. 15.

Smreček R. (2012) - Utilization of ALS data for forestry purposes. GI Forum 2012: Geovizualisation, Society and Learning, Editors: T. Jekel, A. Car, J. Strobl and G. Griesebner, pp. 365-375. ISBN 978-3-87907-521-8.

Sohn G., Dowman I. (2007) - Data fusion of high-resolution satellite imagery and LiDAR data for automatic building extraction. ISPRS Journal of Photogrammetry and Remote Sensing, 62: 43-63. doi: http://dx.doi.org/10.1016/j.isprsjprs.2007.01.001.

Sportouche H., Tupin F., Denise L. (2009) - Building extraction and 3D reconstruction in urban areas from high-resolution optical and SAR imagery. Urban Remote Sensing Event, pp. 1-11.

Stereńczak K., Będkowski K., Weinacker H. (2008) - Accuracy of crown segmentation and
estimation of selected trees and forest stand parameters in order to resolution of used DSM and nDSM models generated from dense small footprint LIDAR data. IAPRS, 37 (B6b): 27-32.

Theng L.B. (2006) - Automatic building extraction from satellite imagery. Engineering Letters, 13: 255-259.

Tian J., Chen D.M. (2007) - Optimization in multiscale segmentation of high resolution satellite images for artificial feature recognition. International Journal of Remote Sensing, 28: 4625-4644. doi: http://dx.doi.org/10.1080/01431160701241746.

Topan H., Maktav D., Jacobsen K., Buyuksalih G. (2009) - Information content of optical satellite images for topographic mapping. International Journal of Remote Sensing, 30: 1819-1827. doi: http://dx.doi.org/10.1080/01431160802642271.

Toutin T. (2004) - Geometric processing of remote sensing images: models, algorithms and methods. International Journal of Remote Sensing, 25: 1893-1924. doi: http://dx.doi.org/10.1080/0143116031000101611.

UTSA (2013) - Object-oriented classification. Lecture 11. http://www.utsa.edu/lrsg/Teaching/EES5083/L11.ppt.

Wiedemann C., Heipke C., Mayer H., Jamet, O. (1998) - Empirical evaluation of automatically extracted road axes. Empirical Evaluation Methods in Computer Vision, IEEE Computer Society Press, Los Alamitos, CA, pp. 172-187.

Yu B., Liu H., Wu J., Hu Y., Zhang L. (2010) - Automated derivation of urban building density information using airborne LiDAR data and object-based method. Landscape and Urban Planning, 98: 210-219. doi: http://dx.doi.org/10.1016/j.landurbplan.2010.08.004.

Zhang Y. (2002) - A new automatic approach for effectively fusing Landsat 7 as well as IKONOS images. Geoscience and Remote Sensing Symposium, IGARSS, pp. 2429-2431.

© 2014 by the authors; licensee Italian Society of Remote Sensing (AIT). This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/4.0/).