An approach for detection of buildings and changes in buildings using orthophotos and point clouds: A case study of Van Erciş earthquake

Gulcan Sarp¹*, Arzu Erener², Sebnem Duzgun³ and Kemal Sahin⁴

¹Department of Geological Engineering, Suleyman Demirel University, 32260 Isparta, Turkey
²Department of Geomatics Engineering, Kocaeli University, 41380 Kocaeli, Turkey
³Department of Mining Engineering Geodetic and Geographic Information Technologies, Middle East Technical University, 06800 Ankara, Turkey
⁴Turksat Directorate of Geographic Information Technologies, Ovecler, Ankara, Turkey

*Corresponding author, e-mail address: gulcansarp@gmail.com

Abstract
This paper presents an image analysis of the Van Erciş earthquake, and demonstrates how efficiently the orthophoto images and point clouds from stereo matching data can be used for automatic detection of buildings and changes in buildings. The proposed method contains three basic steps. The first step is to classify the high-resolution pre and post event Red-Green-Blue (RGB) orthophoto images (orthoRGB) using Support Vector Machine (SVM) classification procedure to extract the building areas. In the second step, normalized Digital Surface Model (nDSM) band derived from point clouds and Digital Terrain Model (DTM) is integrated with the SVM classification (nDSM+orthoRGB). In the last step, building damage assessment is performed through a comparison between two independent classification results from pre- and post-event data. It was observed that using the nDSM band in the classification process as additional bands the accuracy of classification increases significantly.

Keywords: Change detection, earthquake, orthophotos, nDSM, point cloud.

Introduction
Natural disasters are rapid and extreme events that usually result in financial loss, fatality, injury and environmental degradation. In order to minimize the consequences of these catastrophic events, accurate spatial information regarding the geographic extent of the affected areas, both during the outbreak and shortly after the event should be accurately determined. Although earthquakes cannot be prevented, the effective response to them decreases their consequences. Remote Sensing (RS) techniques, both spaceborne and airborne, can have various contributions especially in response and recovery phases of the disaster [Voigt et al., 2007]. The applications of remote sensing science and technologies have been broadly utilized for earthquake damage assessment [e.g. Turker and San, 2003, 2004; Zhang et al., 2003; Stramondo et al., 2006; Sumer and Turker, 2006; Yu et al., 2010;
Various methods have been designed for building damage assessment either to detect changes between pre- and post-event data or to interpret only post-event data. A popular technique to obtain the changes after a disaster is to compare the pre-event Very High Resolution (VHR) imagery with the post-event ones [e.g. Gamba and Casciati, 1998; Turker and San, 2003; Baiocchi et al., 2014a] and the pre-event and post event Digital Surface Model (DSM) data [Baiocchi et al., 2014b]. DSM required for interpretation of building condition; and VHR orthophotos usually provides the one of the best imagery data for damage assessment of buildings. The single use of DSM data or its fusion with other data sources has been the dominant trend of research and practice in feature extraction [Sefercik et al., 2014]. Change detection using both pre- and post-earthquake remote sensing data is another popular method to acquire building damage information [Dong and Shan, 2013]. Change detection approaches, including image enhancement and post-classification comparison methods [Maas and Vosselman, 1999], identify the differences in the state of the building by observing it at different times. The image enhancement method uses arithmetic operations and/or statistical methods to analyze different temporal images such as subtraction of bands, ratioing, and image regression to identify changed areas. Post-classification comparison examines different temporal images after independent classification application. Erener et al. [2012] aimed to monitor the temporal and spatial characteristics of urban development towards 2006 until 2009 by using RS technology integrated with Geographic Information System (GIS). They used high-resolution QuickBird satellite data to inter-compare classified maps for different years quantitatively using change matrices and also qualitatively using spatial change maps. Erener and Yakar [2012] employed Landsat images in order to monitor the dynamics of a lake. They obtained change maps for three different periods: 1987 to 2000, 1987 to 2006 and 2000 to 2006. They investigated coastline changes using these maps. Sarp [2012] used post-classification comparison of Landsat Thematic Mapper (TM) and Landsat Enhanced Thematic Mapper plus (ETM+) data for the years 1989, 2000, and 2010 to detect changes in vegetation.

Most researchers have used multi-class support vector machine (SVM) classification for land use detection of urban areas from high-resolution satellite images. SVMs are powerful tools for providing solutions to classification, regression and density estimation types of problems [Sagale and Kale, 2014]. SVMs are able to model complex nonlinear decision boundaries and are less prone to over fitting than other methods [Adebowale et al., 2013]. It uses a subset of training points in the decision function (called support vectors), so it is also memory efficient. SVMs also have the ability to update the training patterns dynamically whenever there is a new pattern during classification. [Peddabachigari et al., 2004]. Tuia et al. [2010] performed SVM classification using composite kernels for the classification of high-resolution urban images and concluded that a significant increase in the classification accuracy was achieved when the spatial information was used. Bruzzone et al. [2004] proposed a multi-level hierarchical approach to classify high spatial resolution images with SVM, and the results confirmed the effectiveness of the approach. Huang et al. [2008] proposed an adaptive multi-scale information fusion algorithm to extract the spatial features and classify the high-resolution imagery. Li et al. [2010] presented an object-oriented land cover classification method based on SVM. Their results indicate that fusion strategy and classification preprocessing increases classification accuracy.
In this study, the framework for monitoring and assessing heavily impacted areas affected by the Van Erciş earthquake using orthophotos and point cloud data is presented. Integration of the elevation information from point cloud with the spectral information from digital orthophotos eliminates the shortcomings of each data set and takes the advantages of both. By this way, it is possible to improve building extraction accuracy. The SVM classification is performed by using the spectral information as well as the elevation information from point cloud as additional bands. This allows analyst to identify damaged areas in fast and accurate manner, which is essential for effective disaster management. Moreover the effect of using elevation data as additional band on the accuracy of building detection and hence performance of building detection for damage assessments evaluated.

The study area and data sets
The study area is located in Van province in the eastern Turkey. An earthquake (moment magnitude 7.1) occurred on 23 October 2011 at 1:41 pm (local time) with epicenter ~19 km northeast of the Van city [U.S. Geological Survey (USGS), 2011]. The upper left and the lower right coordinates of the study area are 39.033232N - 43.355450E and 39.023632N - 43.370721E, respectively (Fig. 1).

The test region selected for the study is a highly-dense built-up area which involves various urban objects such as collapsed and undamaged buildings with different land uses like roads, bare soil, vegetation and shadow. In some places, collapsed buildings appear indistinguishable from roads, pavements and bare soil and may reflect fragmented characteristics due to shading or they may be occluded by other buildings. Additionally, manmade structures are composed of different sizes and different roof materials such as concrete, brick, asphalt, metal, soil, etc.

The dataset used in the study for the detection of changes caused by the earthquake is composed of DTM, pre and post event orthophotos and point cloud data. Pre event and post event aerial photos were acquired on 7-15 September 2010 and 23 October 2011. Aerial photos were taken from a plane (BeachCraft-B200) using Microsoft Vexcel UltraCam X, S/N UCX-SX-1-40013311 digital aerial camera. The camera has forward motion compensation system. The camera has 9420 pixel × 14430 pixels. The pixel size is 7.2 micron × 7.2 micron. This pixel size corresponds to 16.05 cm ground sampling distance, which provides a high spatial resolution to identify the buildings and collapsed regions. The focal length is 100.5 mm. Inertial Measurement Unit (IMU) system records the rotations around three axes. By aero triangulation process, with the help of ground control and tie points, bundle block adjustment is carried out and exterior orientation parameters of the images are calculated. The aerial photos were acquired approximately 5,910 m high from the geoid in three bands. The ground sampling distance of the images is 30 cm. The photo scale is 1:41,985. They are produced in 8- bit tagged image file format. Approximate projection center coordinates and rotation angles around the coordinate axes measured with the help of a global navigation satellite IMU system. Images were acquired by 60% overlap and 30% side lap. The image dataset covers the whole study area and have zero cloud cover.

The topographic map used in the study has 1:25,000 scale. Topographic maps and pre-event areal images were provided from General Command of Mapping of Turkey and post-event images were obtained from General Directorate of Land Registry and Cadaster of Turkey.
Figure 1 - True color (RGB) composite of the pre-event orthophotos covering the study area. (The aerial photos were acquired approximately 5,910 meters high from the plane).

The Methodology
The method consists of DSM generation by image matching from stereo aerial photos, DTM generation from topographic contour lines, nDSM generation from DTM and DSM, application of SVM classification, filtering the artifacts, accuracy check, and change detection procedures (Fig. 2). Each step is defined as follows:

Figure 2 - Flow chart of the proposed methodology.
**DSM generation by image matching from stereo aerial photos**
The point cloud data includes heights computed from stereo image pairs using image matching techniques. This technique has been successfully used to extract DSM from stereo image pairs. It is composed of the process of finding matching points in the corresponding stereo image pairs. The conjugate point in the corresponding image is determined by comparing the certain characteristics of the entity on the basis of similarity measure, generally the cost function like correlation coefficient. Theoretically, the main task in image matching is to select a matching point in one image, finding its conjugate point in the other image (Stereomate), computing 3D position of the matched points in object space, and finally, assess the quality of the matching [Schenk, 1999].
The urban environment is always considered as a challenge for accurate extraction of the elevation data. To extract DSM from VHRS (Very-High-Resolution-Satellite) images, the perspective projection of the image, occlusion, shadow, surface discontinuities such as buildings and trees, abrupt change in height, large areas with little or no texture, and representative similar patterns should be taken into consideration [Li and Gruen, 2004].

**DTM generation from topographic contour lines**
DTM is an elevation model, which depicts the pure terrain surface without buildings and vegetation. Hence, elevation is given in topographic maps. In this study, DTM data were generated from 1:25,000 scale topographic maps with a cell size of 10 m after creation of Triangular Irregular Network (TIN) model. In order to test the accuracy of the produced DTM, the principles of DTM accuracy assessment given by USGS were adopted. According to USGS National Mapping Program Technical Instructions, a representative sampling of test points was used to verify the accuracy of any category of the DTM. A minimum number of 28 test points for DTM is required [USGS 1998]. The root-mean-square error (RMSE) statistic for topographical elevation was used to describe the vertical accuracy of the DTM. The RMSE of the result data is 1.02 m. This error is quite admissible within the accuracy requirements of the study, which is much lesser than the minimum accuracy limit of one-third of the contour interval.

**Normalized DSM (nDSM) Generation**
In the building detection procedure, the basic idea of using an nDSM is that the man-made objects with different heights above the terrain can be detected by applying a threshold to nDSM. The various heights of terrain will affect the building extraction, especially when height threshold cannot be avoided. Thus, the DTM and the nDSM are necessary for building extraction. The DSM includes the objects with their heights above the ground as well as the topography. To calculate nDSM, DTM was subtracted from the DSM.

**Support Vector Machines (SVM) Classification**
SVM is a supervised learning system and is based on recent advances in statistical learning theory [Cristianini and Shawe-Taylor, 2000]. Cortes and Vapnik [1995] developed SVM for binary classification. There are a number of publications for the mathematical formulation of the SVM [e.g. Vapnik 1995, 1998; Burges 1998].
SVM separates the classes with a decision surface that maximizes the margin between the classes. The surface is often called the optimal hyperplane, and the data points, which are
closest to the hyperplane, are called support vectors (Fig. 3). The support vectors are the critical elements of the training set [Boser et al., 1992; Cortes and Vapnik 1995; Foody et al., 2007].

SVM need training data that optimize the separation of the classes rather than describing the classes themselves [Foody and Mathur, 2006]. Using a Radial Basis Function (RBF), class distributions with non-linear boundaries can be mapped into a high dimensional space for linear separation [Huang et al., 2002]. Training the SVM with a Gaussian RBF requires setting two parameters: regularization parameter that controls the trade-off between maximizing the margin and minimizing the training error, and kernel width. A small regularization parameter tends to emphasize the margin while ignoring the outliers in the training data, while a large regularization parameter may overfit the training data. A comprehensive description of parameters of SVM can be found in Burges [1998] and Cristianini and Shawe-Taylor [2000].

SVM classifier provides four types of kernels: linear, polynomial, RBF, and sigmoid. The RBF kernel works well in most cases [ENVI Manual, 2004]. The mathematical representation of each kernel is listed in equations [1-4],

\[
\text{Linear} : K(x_i, x_j) = x_i^t x_j \quad [1]
\]
Polynomial: $K(x_i, x_j) = (\gamma x_i^t x_j + r)^d, \gamma > 0 \quad [2]$

RBF: $K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2), \gamma > 0 \quad [3]$

Sigmoid: $K(x_i, x_j) = \tanh(\gamma x_i^t x_j + r) \quad [4]$

where $x_i$ is $i^{th}$ support vector and $x_j$ is $j^{th}$ training data points, $t$ is the smoothing parameter, $K$ is the kernel function, $\| \|$ is the Euclidean norm, $\gamma$ is kernel width in the kernel function for all kernel types except linear, $d$ is the polynomial degree term in the kernel function for the polynomial kernel, $r$ is the bias term in the kernel function for the polynomial and sigmoid kernels, $\gamma$, $d$, and $r$ are user controlled parameters, as their correct definition significantly increases the accuracy of the SVM solution.

Filtering the artifacts
Morphological operations have been proved being capable of deleting small disconnected regions, filling cavities, and smoothing the region-of-interest [Giardina and Dougherty, 1988]. These operations rely only on the relative ordering of pixel values, not on their numerical values, and therefore are especially suited to the processing of binary images. Morphological operations, opening followed by closing, are applied to the image to eliminate the small isolated regions. Therefore, the global shape of the objects was not distorted [Sonka et al., 1998]. These operations are defined as ordered combinations of fundamental operations, dilation, and erosion. Morphological operators of opening and closing were applied after performing the SVM classification. Due to misclassification, the classified building areas contained artefacts. Therefore, these artefacts were removed using the morphological operations. An opening filter removes thin protrusions, outward pointing boundary irregularities, thin joins, and small isolated objects. On the other hand, a closing filter removes the thin gulf, the inward-pointing boundary irregularities and small holes [Gonzalez et al., 2004]. Therefore, combinations of closing and opening operations provide an effective result for removing the artefacts and noise [Aytekin et al., 2012]. After SVM classification, non-building areas were masked out from the resultant images, since the study focuses on building features. The misclassified artifacts were removed from the building areas by the help of the morphological filters. In this case, closing followed by the opening was applied to the building areas.

Accuracy assessment
In order to assess the accuracy of the proposed methodology, the boundaries of buildings were digitized from 2010 and 2011 images and then labeled in the GIS environment. The building boundaries used as ground truth data were compared with the resultant image obtained using the above methodology. The ground truth data and the output map of the
Building change detection

method were analyzed in the Matlab environment. The accuracy assessment involves computation of True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) components by the comparison of ground truth and the resultant image. TP refers to the regions determined as a feature (building or road) both in the ground truth and in the resultant image. TN refers to the regions that are not determined as a feature in the ground truth and in the resultant image, respectively. FN refers to the features that cannot be determined and FP refers to the regions, which are determined as features although these are not present in the ground truth [Shufelt and Mckeown, 1993]. Based on these components, the split factor (SF), missing factor (MF), percent of building detection (PBD), and quality percent (QP) are calculated using the following equations [5-8],

\[
SF = \frac{FP}{(TP + FP)} \quad [5]
\]

\[
MF = \frac{FN}{(TP + FP)} \quad [6]
\]

\[
PBD = 100 \times \frac{TP}{(TP + FP)} \quad [7]
\]

\[
QP = 100 \times \frac{TP}{(TP + FP + FN)} \quad [8]
\]

Post classification comparison

Building damage was detected through making a comparison between two independent classification results from pre and post event data. The main advantage of these types of methods is minimizing the effect of radiometric difference between the two data sets [Coppin et al., 2004]. However, the accuracy is mainly dependent on the initial classification results of the SVM classification.

The change detection between pre-event and post-event data was performed by comparison of building objects in the GIS environment. The pre-event and post-event buildings were matched and compared by using an overlay technique. In order to do the operation, the buildings obtained by SVM algorithm from 2010-2011 Red-Green-Blue (RGB) orthophoto images (orthoRGB) and 2010-2011 nDSM+orthoRGB were converted into vector format to test the matching segments of the buildings. This operation produces three types of output. The first output was matching buildings of 2010 and 2011 nDSM+orthoRGB representing buildings of no-change, the second output was non-matching buildings of 2010 and 2011 nDSM+orthoRGB representing collapsed buildings and the third output was non-matching buildings of 2011 and 2010 representing newly build up buildings.
**Results and discussions**

To derive a detailed DSM, an interpolation was done for matched points. First, a point cloud was interpolated into a grid format. The nearest neighbor interpolation method was employed to preserve the sharp leap in elevation along the edges of buildings. The cell size was chosen as close as possible to the point density. The grid value then represents the elevation. The accuracy of each grid depends on the digital camera used. The digital camera provides a horizontal accuracy of 5.35 cm and a vertical accuracy of 19.8 cm. The horizontal accuracy is similar to that of an analog camera with similar frame length whereas vertical accuracy is lower as compared to that of the analog camera. Figure 4a and 4b provide the DSMs for 2010 and 2011 generated from the point cloud data. The horizontal and vertical accuracy values of the images obtained from the digital camera used in this study may not be reached for the images obtained from the VHR satellite sensors, like Ikonos, Quickbird, Geoeye, Worldview-2. As such data was not available for the site the comparison with satellite data is not considered in this study.

nDSM for the years of 2010 and 2011 are given in the Figure 4c and 4d respectively. In an nDSM the streets have gray level values near zero. The aboveground objects have gray level values according to their real heights. In these binary images, buildings can be roughly separated according to the height information. However, obtaining a clear boundary is hard to achieve due to the high building density. The most severe problem was to distinguish between building and vegetation areas. The size criterion for the selection was not sufficient for larger vegetation areas or vegetation areas close to buildings. Moreover, direct height subtraction would not be sufficient, due to low contrast in building representation. As the images used in this study are aerial photos with RGB bands, the vegetation indices, which involve NIR and R bands, cannot be used to identify vegetation areas. In this respect VHR multispectral satellite images, like Ikonos, Quickbird, Geoeye, Worldview-2, would have advantages to distinguish vegetation areas.

The DTM and DSM used in this study have RSME of 1.19 and 1.02 m, respectively. Such errors decrease the building detection accuracy especially for the built-ups with clusters of adjacent buildings, where nDSM has combined error of DSM and DTM and built-ups with relatively short buildings surrounded by tall trees or tall buildings, where three heights or tall buildings are not distinguished in nDSM.

In this study, the orthophoto images of 2010 and 2011 were classified into the five categories as building, road, vegetation, shadow, and bare land through SVM classification. The classification training samples were collected from the representative homogeneous areas. The RBF was selected as the kernel method for the SVM classification. This function can handle linearly non-separable problems and works well in most cases [ENVI Manual, 2004]. $\gamma$ was determined as the inverse of the number of bands in the input image and 1000 was taken for the value of the regularization parameter. The results of the classified RGB image, and nDSM+orthoRGB representing the classes for the 2010 and 2011 are given in the Figure 5a, 5c, and Figure 5b, 5d, respectively.
Figure 4 - Image representing the DSM and nDSM aboveground features (a) DSM of 2010, (b) DSM of 2011, (c) nDSM of 2010, and (d) nDSM of 2011.

Figure 5 - (a) SVM results of 2010 orthoRGB, (b) SVM results of 2010 nDSM+orthoRGB; (c) SVM results of 2011 orthoRGB; (d) SVM results of 2011 nDSM+orthoRGB.
The pixel-based performance evaluation results are given in Table 1. The results of the accuracy assessment showed that PBD for 2010 orthoRGB and 2010 nDSM+orthoRGB were 86.10% and 90.39%, respectively and PBD for 2011 orthoRGB and 2011 nDSM+orthoRGB were 83.95% and 89.76%, respectively. The QP change between 41.82% and 50.10% for 2010 orthoRGB and 2010 nDSM+orthoRGB and 45.80% and 52.34% for 2011 orthoRGB and 2011 nDSM+orthoRGB composite images, respectively. The results of the SVM classification of the RGB and nDSM integrated bands of pre and post event images provide more promising results than SVM classifications of the RGB bands since the integration of the elevation information from nDSM with the spectral data from digital orthophotos provides more information.

<table>
<thead>
<tr>
<th>Data Set</th>
<th>Number of Pixels</th>
<th>Ratio</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TP</td>
<td>FP</td>
<td>FN</td>
</tr>
<tr>
<td>SVM of 2010 orthoRGB</td>
<td>10230</td>
<td>12579</td>
<td>1651</td>
</tr>
<tr>
<td>SVM of 2010 nDSM+orthoRGB</td>
<td>10739</td>
<td>9555</td>
<td>1142</td>
</tr>
<tr>
<td>SVM of 2011 orthoRGB</td>
<td>9706</td>
<td>9632</td>
<td>1856</td>
</tr>
<tr>
<td>SVM of 2011 nDSM+orthoRGB</td>
<td>10378</td>
<td>8267</td>
<td>1184</td>
</tr>
</tbody>
</table>

True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN), Split Factor (SF), Missing Factor (MF), Percent of Building Detection (PBD), Quality Percent (QP), Support Vector Machine (SVM), Red-Green-Blue (RGB), normalized Digital Surface Model (nDSM).

The results of the overlay analysis indicates that the use of nDSM data as an ancillary band in the SVM classification of the RGB bands of the orthoRGB would be useful to better distinguish adjacent buildings from each other and to improve the shape of the detected buildings. The analysis indicates that many of the no-change (Fig. 6b - Case1), collapsed (Fig. 6b - Case2) and newly built-up buildings (Fig. 6b - Case3) were easily detected, and generated results were visually compared with the orthoRGB data. Depending on these results, 1250 matching buildings, 300 collapsed buildings and 173 newly build buildings were detected in the study area (Fig. 6a). The overall results of the change data can be considered relatively compatible with the results of the approach suggested in this research.
Figure 6 - (a) Building change detection results of 2010 and 2011 nDSM+orthoRGB SVM classification. (b) Examples of different change classes. Case 1: no-change buildings; Case 2: collapsed buildings; Case 3: newly build buildings.
Conclusion
Damage assessment studies related to post earthquake scenario involves various kinds of challenges for identification and characterization of collapsed and undamaged buildings. Damaged urban regions contain collapsed buildings, which may appear indistinguishable from roads, pavements, and bare soil and may be occluded by other buildings. Thus, due to the complexity of the problem, it is hard to extract and delineate undamaged and collapsed buildings. In order to increase the accuracy of building extraction for damage assessment, the elevation information to SVM classification as an additional band is added. The results of the study indicate that the inclusion of the nDSM to SVM classification increased the classification accuracy for building detection hence increase the performance for damage assessment.

The approach provided in this study can be adopted for satellite sensors, which provide VHR stereo images, like Ikonos, Quickbird, Woldview-2. Although the spatial resolution of these satellite images are lower than the aerial photos, the additional spectral information like NIR and other bands, allow analyst to better separate vegetation and building areas.

Once the images are obtained, the damage assessment based on the proposed approach takes almost 36-48 hours, which have sufficient potential to support disaster response activities. Moreover, such a damage assessment approach is much more timesaving than ground-based damage assessment, which is required for post disaster planning to be performed by the insurance companies, which are supposed to evaluate the losses and payouts, local and global disaster managers who are responsible for selecting appropriate shelters sites, allocating services for affected population.

For the damage assessment due to other natural disasters like landslides, floods, tsunamis, the proposed approach needs to be adapted based on the nature of the hazard. For example, in case of landslides the change in the terrain and hence the DSM is much more pronounced than the case of earthquake and hence the nDSM threshold should be adjusted accordingly. Similarly, as the floods and tsunamis involve water intrusion into built-up areas, use of spectral bands reflecting water areas as well as the changes in the nDSM due to water intrusion should be taken into account. The authors’ further research will focus on these adaptations.

Acknowledgment
We would like to thank the anonymous reviewers for their constructive comments.

References


Sumer E., Turker M. (2006) - *An integrated earthquake damage detection system*. In: The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Salzburg, Austria, XXXVI, 4/C42.


