

## PERFORMANCE ASSESSMENT OF FULLY AUTOMATIC THREE-DIMENSIONAL CITY MODEL RECONSTRUCTION METHODS

Beril Sirmacek, Hossein Arefi, Thomas Krauss

German Aerospace Center (DLR), Earth Observation Center (EOC)  
PO Box 1116, 82234, Wessling, Germany  
(Beril.Sirmacek, Hossein.Arefi, Thomas.Krauss)@dlr.de

Commission III/4

**KEY WORDS:** Building, Modelling, City, DSM/DTM

### ABSTRACT:

Three-dimensional urban region representations can be used for detailed urban monitoring, change and damage detection purposes. In order to obtain three-dimensional representation, one of the easiest and cheapest way is to use Digital Surface Models (DSMs) which are generated from very high resolution stereo satellite images using stereovision techniques. Unfortunately after applying the DSM generation process, we cannot directly obtain three-dimensional urban region representation. In the DSM which is generated using only one stereo image pair, generally noise, matching errors, and uncertainty on building wall locations are very high. These undesirable effects prevents a DSM to provide a realistic three-dimensional city representation. Therefore, some automatic techniques should be applied to obtain three-dimensional city models using DSMs as input. In order to solve the existing problems in this field, herein we introduce two automated approaches based on usage of DSMs as input. The first method depends on using of a 3D active shape model for building shape extraction and 3D reconstruction, the second approach is based on an approximation of prismatic models to DSMs. Our experimental results on images and DSMs of Tunis city which are obtained from WorldView-2 satellite indicate possible usage of the proposed algorithms to obtain three-dimensional city representations automatically.

### 1 INTRODUCTION

Three-dimensional representations of complex environments obtained a lot of interest for various applications with the development of satellite sensor technology. In order to obtain 3D representations of cities, one of the most practical method is to generate Digital Surface Models (DSMs) using very high resolution satellite images which are taken from two or more viewing angles. Unfortunately, due to occlusions, matching errors, and applied interpolation techniques these DSMs do not represent 3D city models directly. Automatically generated DSMs do not contain sharp wall and rooftop representations. Therefore, in order to obtain real 3D city representations, advanced methods should be applied to DSMs.

In the related literature, Brunn and Weidner (Brunn and Weidner, 1997) used surface normals on DSM to discriminate buildings and vegetation. After detecting buildings, they measured the geometry of rooftops using surface normals and they interpolated polyhedral building descriptions to these structures. Fradkin et al. (Fradkin et al., 1999) proposed a segmentation based method to reconstruct three-dimensional models of dense urban areas. To this end, they used very high resolution color aerial images and disparity maps. Canu et al. (Canu et al., 1996) used a high resolution DSM, which is obtained by stereo matching techniques, in order to reconstruct three-dimensional buildings. First, they segmented the DSM into homogeneous regions. Then, they interpolated flat surfaces on these regions. Ortner et al. (Ortner et al., 2002) used a point process (Jacobsen, 2005) to model urban areas. They represented urban areas as interacting particles where each particle stands for an urban object. Preknowledge about building shapes is used to model these particles. Arefi et al. (Arefi et al., 2008) extracted above-ground objects from LIDAR data. Then, three-dimensional buildings are reconstructed by hierarchical fitting of minimum boundary rectangles (MBR) and a RANSAC based straight line fitting algorithm. Tournaire

et al. (Tournaire et al., 2010), developed a stochastic geometry based algorithm to detect building footprints from DSM data which have less than 1 m resolution. They tried to fit rectangles on the buildings using an energy function and prior knowledge about buildings. To minimize the energy function, they used a Reversible Jump Monte Carlo Markov Chain (RJMCMC) sampler coupled with a simulated annealing algorithm which leads to an optimal configuration of objects. Maas (Maas, 1999) used maximum slope values in order to determine best fitting rooftop shapes to generate three-dimensional building models. Valero et al. (Valero et al., 2008) developed a feature extraction and classification based method to classify building roofs into two classes as flat-roof and gable-roof. They estimated ridge-line positions which are based on skeletons of groundfloor plans. They provided the difference between the average roof outline height and the average ridge-line height as first feature, and the norm of the orthorectified image gradient as second feature for the support vector machine (SVM) classifier. In all introduced studies, good results are achieved generally using very high resolution (better than 1 m spatial resolution) DSMs which are generally generated from airborne images or LASER scan data. However, enhancement of buildings in low resolution urban DSM data which are generated from satellite images is still an open research problem. In order to bring an automated solution to this problem, in previous work Sirmacek et al. proposed a novel technique for obtaining three-dimensional city representations by applying a building shape and rooftop-type detection approach to DSMs (Sirmacek et al., 2011). They started by applying local thresholding to raw DSMs in order to extract high urban objects which can indicate building locations. They extracted building shapes from regions which are obtained from thresholding result by using a binary active shape growing algorithm that they proposed which depends on growing rectangular shapes in elongated segments which are detected in binary mask which is obtained by thresholding the DSM. After obtaining building shapes, they generated three-dimensional models by understanding building

rooftop-type.

Herein, we introduce two different fully automatic techniques to generate 3D city models from DSMs. Both methods need a DSM and a segmentation result as inputs. The input DSM is generated from stereo satellite images by using the semi-global matching (SGM) approach based on combination of census and mutual information. The segmentation result is obtained automatically from the DSM by deriving a terrain model (DTM) and fusing the derived high object mask with vegetation and water masks derived from the pansharpened multispectral ortho images fitting on the DSM.

The first 3D modeling approach consists of three main modules as: detection of building shapes, rooftop classification, and 3D reconstruction based on the detected shapes and rooftop classes. This approach extracts seed-point locations from the segmentation result which indicate approximate building regions. A novel active rectangular shape growing approach is used on these seed-point locations to approximate the building shapes on the DSM. Afterwards, using derivative filters on DSM, we extract roof ridge-lines automatically. Obtained ridge-line information helps to classify building rooftops as flat or gable shapes. Finally, regarding the rooftop class we assign rooftop models to our 3D model.

The second 3D reconstruction approach is based on hierarchical approximation and merging the segmented regions providing prismatic models of the buildings according to LOD1 of CityGML standard. Building footprints are approximated into regular polygons by reducing the boundary pixels into the most significant nodes which preserve the shape and size of the original segment. The building shapes are first classified into rectilinear and non-rectilinear by measuring the orientation of the edges. For a rectangular building containing one main orientation of the edges a method based on Minimum Bounding Rectangle (MBR) is employed. In contrast, a Combined Minimum Bounding Rectangle (CMBR) approach is proposed for regularization of non-rectilinear polygons, i.e. buildings with not perpendicular edge directions. Both MBR- and CMBR-based approaches are iteratively employed on building segments to reduce the original building footprints to a minimum number of nodes with maximum similarity to original shapes. Finally, prismatic models (LOD1) are created by computing an average height of the internal building pixels and generating the roof and wall polygons.

## 2 OBTAINING BUILDING SEGMENTS

For the extraction of building segments first a spectral classification using a rule based fuzzy method and second a derivation of a digital terrain model (DTM) from the digital surface model (DSM) delivered from the semi global matching algorithm (SGM) is derived.

The rule based fuzzy classification uses the two fuzzy functions

$$\text{Fuzzy up defined as } f^+_{c,d} x = \begin{cases} 0 & \text{if } x < c \\ \frac{x-c}{d-c} & \text{if } c \leq x \leq d \text{ and Fuzzy} \\ 1 & \text{if } x > d \end{cases}$$

down  $f^-_{a,b} = 1 - f^+_{a,b}$  together with the logical functions *Fuzzy*

and  $\mathcal{F}(a, b, \dots) = \min(a, b, \dots)$  and *Fuzzy or*  $\mathcal{G}(a, b, \dots) = \max(a, b, \dots)$ . The rules used for the classification are based on the channels of WorldView-2 (coastal, blue, green, yellow, red, red-edge, near infrared, second near infrared) denoted as  $(C, B, G, Y, R, E, N, S)$  and the three ratio functions  $ndvi = \frac{N-R}{N+R}$ ,  $cni = \frac{N-C}{N+C}$  and  $bgi = \frac{B-Y}{B+Y}$ . Using these and a fuzzy range  $d = 2\%$  the classes vegetation ( $v$ ), soil ( $s$ ), water ( $w$ )



Figure 1: Top-Of-Atmosphere (TOA) reflections image showing a section 1 km  $\times$  1 km from a WorldView-2 triple-stereo-scene acquired 2011-04-13 over Tunis, channels 2, 3 and 5 (blue, green, red)

and shadow ( $\sigma$ ) are defined on the previously calculated top-of-atmosphere (TOA) reflectances image in [%] (Fig. 1) as:

$$v = f^+_{0,0.2} ndvi \quad (1)$$

$$s = \mathcal{F} \left( f^-_{-d,d} B - G, f^-_{-d,d} G - Y, f^-_{-d,d} Y - R, f^-_{-d,d} R - N \right) \quad (2)$$

$$w = \mathcal{F} \left( f^+_{-d,d} B - G, f^+_{-d,d} G - Y, f^+_{-d,d} Y - R, f^+_{-d,d} R - N, f^+_{-d,d} B - 7.5\% \right) \quad (3)$$

$$\sigma = \mathcal{F} \left( f^+_{0.078,0.39} ndvi - cni, f^-_{-d,0} B - 5\% \right) \quad (4)$$

Obtaining from these fuzzy results ranging from 0 to 1 the binary masks shown in Fig. 2 the limits  $w > 0.59$ ,  $v > 0.89$ ,  $s > 0.59$  and  $\sigma > 0.25$  are applied.

In the second step a digital terrain model (DTM) is derived from the generated digital surface model (DSM). For this the DSM is scaled down by an factor of eight filling already small unmatched (no-data) regions. Afterwards a morphological grayscale opening using a structuring element of radius 20 (i.e.  $20 \times 8 \times 0.5$  m) with a 10 % or 90 % percentile respectively is used filling also remaining no-data areas. Applying the final rescaling to original size gives the filled DTM as shown in Fig. 3. Subtracting the DTM from the DSM provides us with the so called normalized digital elevation model (nDEM) which gives the heights of objects above ground. A profile showing the connection between DSM, DTM and nDEM is shown in Fig. 3, right. The red marked objects are the building segments used in the following steps.

Applying a threshold filter at the appropriate height (here: 3 m) and a morphological opening followed by a closing (radius of

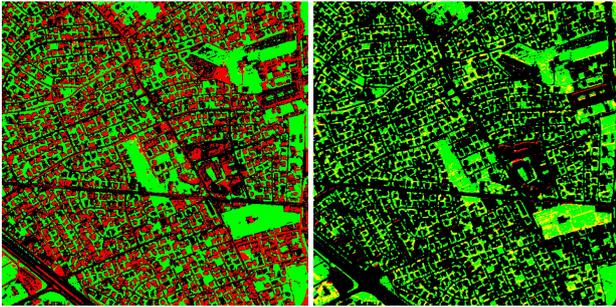


Figure 2: Classifications derived from pan sharpened TOA image, left: water (blue), vegetation (green), sealed/soil (red), dry vegetation (yellow), right: water (blue), vegetation (green), shadow (red), shadow+vegetation (yellow, mostly near trees)

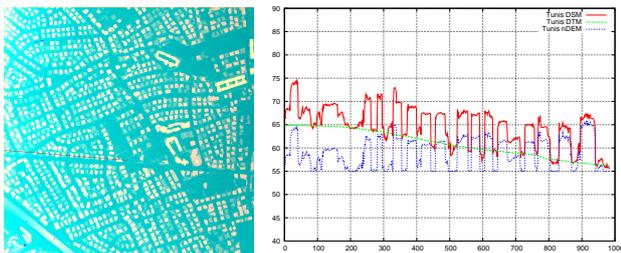


Figure 3: Left: Overlay of DSM (green), DTM (blue) and nDEM (red) of Tunis area with depicted profile line (red); right: profiles along profile line: DSM (green), DTM (blue) and nDEM (red) (heights in meter above WGS84 ellipsoid)

structuring element: 2 px) gives the *high objects* from the nDEM (cf. Fig. 4, left). Combining the classification and these high objects mask gives the classes buildings (high and non-vegetation), trees (high and vegetation), ground (all non high objects including water, vegetation and non-vegetation areas) as shown in Fig. 4, right.

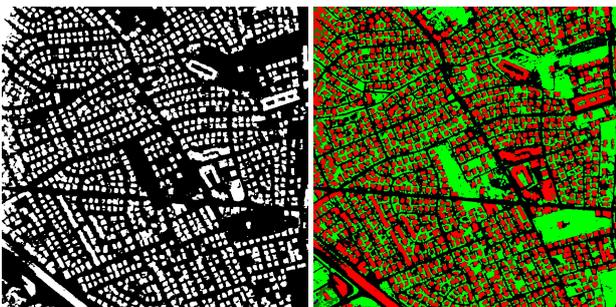


Figure 4: Left: High objects mask, right: final classification, red: buildings, green: low vegetation, dark green: trees, blue: water

### 3 THREE-DIMENSIONAL CITY MODEL RECONSTRUCTION

In this section, we introduce two different approaches for three-dimensional city model reconstruction. In the experiments section, we compare these two different approaches and discuss about their detection capabilities, advantages and disadvantages.

#### 3.1 Active Shape Growing Based Approach

The active shape growing approach is our first 3D city model reconstruction approach. This approach is based on three main modules as: active shape growing for shape detection, detecting building rooftop type, and finally we use detected the shapes and rooftop types to generate 3D models.

**3.1.1 Box-Fitting for Shape Detection** For detecting complex building shapes, herein we follow a similar methodology as represented in the previous study (Sirmacek et al., 2011). However instead of using a binary active shape growing approach in each seed-point location, we propose a novel active shape growing approach based on usage of three-dimensional information. To do so, after extracting  $(x_s, y_s)$  seed point locations as we describe in (Sirmacek et al., 2011) in detail, we start to grow our active rectangular shapes in each seed point location by regarding the height information. We assume that  $(x_v^n, y_v^n)$  array holds the pixel coordinates for  $n$ th edge of the virtual rectangular shape. Iteratively, we sweep each edge to the outwards growing direction if the edge pixels satisfy  $(\max(D(x_v^n, y_v^n)) - \min(x_v^n, y_v^n) < \delta)$  inequality ( $n \in [1, 2, 3, 4]$ ). Here, the  $\delta$  threshold is the minimum building height that we would like to detect in the region.

In our application we assume  $\delta$  as equal to 3, which means that we assume the buildings to be higher than 3 meters to be detected. When the growing process stops for each edge, we calculate the final energy value by using the equation that we represent in Eqn. 5. In the equation,  $m(\cdot)$  represents the mean value. For the same seed-point, we apply growing process for all  $\theta \in [0, \pi/6, \pi/3, \pi/2, 2\pi/3, \dots, 2\pi]$  angles with  $\theta_{dif} = \pi/6$  radian turning steps. As we discussed in detail in (Sirmacek et al., 2011), by reducing  $\theta_{dif}$  step sizes, we can obtain more accurate approximations, however in this case we need more computation time. After calculating  $E_\theta$  for all  $\theta$  angles we pick the estimated box which shows the highest  $E_\theta$  energy as detected building shape. Since most buildings appear like composition of rectangular building segments, it makes sense to extract rectangular shapes on buildings. The main advantage of using the box-fitting approach is that approximate building shapes still can be found even if the building edges are not well-determined, or even if there are trees adjacent to the building facades. However, other region growing algorithms fail to extract an object shape in these cases, since the growing region can flow out easily when the parameters are not set precisely.

$$E_\theta = 4 \times m(B_D(x, y) \times D_i(x, y)) - \sum_{n=1}^4 m(D(x_v^n, y_v^n)) \quad (5)$$

For complex buildings, after fitting a chain of boxes, discontinuity between adjacent boxes should be smoothed. For this purpose, we simply benefit from morphological operations. First, we start with filling inside of the detected binary boxes with 1 value in  $B_B(x, y)$  binary mask. Then, we apply morphological dilation and erosion operations respectively to the detected boxes, using a disk shaped structuring element with radius 1. After this operation small discontinuity between adjacent boxes can be smoothed. Final improved building shapes are kept in new  $B(x, y)$  binary mask.

**3.1.2 Classification of Building Rooftops** After detecting building groundfloor shapes, we focus on reconstruction of rooftops. For this purpose, we benefit from our previous ridge-line detection approach which is presented in (Sirmacek et al., 2011).

The ridge-line detection approach is based on derivative calculation over the DSM. We use the following derivative filter. For a symmetric Gaussian function  $G(x, y) = \exp(-(x^2 + y^2))$ , it is possible to define basis filters  $G_{p0}$  and  $G_{p\frac{\pi}{2}}$  as follows,

$$G_{p0} = \frac{\partial}{\partial x} G(x, y) = -2x \exp(-(x^2 + y^2)) \quad (6)$$

$$G_{p\frac{\pi}{2}} = \frac{\partial}{\partial y} G(x, y) = -2y \exp(-(x^2 + y^2)) \quad (7)$$

We can find a derivative in an arbitrary  $\theta$  direction using following filter function,  $G_{p\theta} = \cos(\theta)G_{p0} + \sin(\theta)G_{p\frac{\pi}{2}}$

We convolve our DSM with this derivative filter in  $\theta \in [0, \pi/12, \dots, 23\pi/12]$  directions as follows,  $J_\theta = D(x, y) * G_{p\theta}$

If  $\theta$  angle is perpendicular to the building ridge-line orientation, then one side of the building rooftop gives positive response, and the other side of the building rooftop gives negative response to filter. Assuming  $B^j(x, y)$  is the  $j$ th connected component in  $B(x, y)$  binary building segment matrix, we assume that two sides of the building rooftop ( $RP_\theta^j$  and  $RN_\theta^j$ ) can be extracted as follows,

$$RP_\theta^j = B^j(x, y) \times (J_\theta > 0) \quad (8)$$

$$RN_\theta^j = B^j(x, y) \times (J_\theta < 0) \quad (9)$$

Since we have no pre-information about building orientations, we should do the derivative filtering in all possible orientations. Therefore, we calculate  $\sum_\theta RL_\theta^j$  for  $\theta \in [0, \pi/12, \dots, 23\pi/12]$  directions. Building ridge-line will be extracted in  $\theta_j$  filtering angle which is almost perpendicular to the ridge-line orientation. However the ridge-line will be also detected in  $\theta_j - \pi/12$  and  $\theta_j + \pi/12$  neighbor filtering directions. Therefore, the ridge-line will have a value of higher than 2 in the  $\sum_\theta RL_\theta^j$  result. As a result, ridge-line of  $j$ th building rooftop can be obtained by calculating  $R^j(x, y) = \sum_\theta RL_\theta^j > 2$  and eliminating connected components which are less than 10 pixels in order to eliminate redundant information coming from small objects on rooftop.

Next, we use detected ridge-lines for roof-type classification and three-dimensional rooftop model reconstruction purposes. We benefit from detected roof ridge-lines to classify rooftops as 'flat roof' or 'gable roof' type. If there is no ridge-line detection result on a building segment, we assume that building as flat roof. If there is a ridge-line on the  $j$ th building segment, then we calculate mean of DSM height values on ridge-line location by calculating  $\sum_{(x,y)} R^j(x, y) \times D(x, y) / M$ , where  $M$  is the total number

of ridge-line pixels in  $R^j(x, y)$  binary matrix. We also calculate mean of DSM height values on building border by calculating  $\sum_{(x,y)} B^j(x, y) \times D(x, y) / N$ , where  $N$  holds the total number of building border pixels in  $B^j(x, y)$  binary matrix. If the difference between these two mean values is lower than 2 meters, then we assume the rooftop as a *flat rooftop*. Otherwise, it is assumed as *gable rooftop*. This 2 meters criteria is obtained by observing *gable rooftop* characteristics over test area.

**3.1.3 Obtaining Three-Dimensional Representation** We start with generating a zero matrix  $D_2(x, y)$  with the same size with  $D(x, y)$  matrix. New height values belonging to objects in the city will be stored in  $D_2(x, y)$  matrix. In order to eliminate noise in non-built regions, we apply median filtering to the original DSM ( $D(x, y)$ ) by using a  $3 \times 3$  pixel size window obtaining a filtered DSM ( $D_f(x, y)$ ). For non-builtup areas, or in other areas for  $(x, y)$  coordinates which satisfy  $B(x, y) = 0$ , we apply  $D_2(x, y) = D_f(x, y)$ . That means: we assign smoothed ground height values for non-built regions. As building wall, we insert single height value to each building boundary which is

stored in  $B(x, y)$  binary building shape matrix. For each building, the building height value is assigned by calculating the mean of  $D(x, y)$  values on all building boundary pixels. After smoothing the noise on the ground and inserting sharp building walls, finally rooftop height values are assigned to  $D_2(x, y)$  matrix. Rooftop height assignment is done by considering the roof-type. If the roof is classified as a flat roof, then only a single height value is assigned to all building area which is equal to building wall heights. If the rooftop is classified as a gable-roof, we follow different approaches depending on the building complexity. If the building is not detected as a complex building as structure type (if the Euler number is equal to 1), then we can find polygons which defines the rooftop in three-dimensional space. To this end, we detect corners of building segment using Harris corner detection algorithm (Harris and Stephens, 1988) over panchromatic image of the test region. Besides, we also detect endpoints of the building ridge-line. We pick each building corner one by one. Then, we find the closest ridge-line endpoint. A line between building corner and the closest ridge-line endpoint can divide the rooftop into polygons. A detailed demonstration of this approach is illustrated in (Sirmacek et al., 2011). Height values of rooftop polygons are assigned to corresponding pixels in  $D_2(x, y)$  matrix. If the building is detected as a complex structure or if the building ridge-line could not be extracted properly, unfortunately we cannot use the same idea for building rooftop reconstruction and more advanced rooftop reconstruction approaches are needed. Herein we leave the building rooftop reconstruction at this level. For complex building rooftops and for rooftops for which we cannot extract the ridge-line properly, we only insert corresponding pixel values from  $D_f(x, y)$  matrix.

### 3.2 Hierarchical Generation of 3D Prismatic Models of Buildings

An approach is presented to extract 3D model of buildings according to LOD1 (Level-of-Detail 1) based on CityGML standard (Kolbe et al., 2005). The approach starts with approximation and regularization of the building boundaries. Two algorithms are proposed for approximation of the building polygons based on the main orientation of the buildings (Arefi et al., 2007). The algorithms are selected according to the number of main orientations of the buildings and implemented as follows:

- If the building is formed by a rectilinear polygon, i.e., sides are perpendicular to each others from the top view, a method based on Minimum Bounding Rectangle (MBR) is applied for approximation. This method is a top-down, and model-based approach that hierarchically optimizes the initial rectilinear model by fitting MBR to all details of the data set. Principles of MBR based polygon approximation is presented in Figure 5.

Accordingly, after determination of the main orientation, the building polygon is rotated to the main orientation, as shown in Figure 5(a). In the next step the MBR image (Fig. 5(b)) is subtracted from the rotated building region. After subtraction, new regions will be produced (Fig. 5(c)). For any of those regions a MBR will be calculated (Fig. 5(d)). They are again subtracted from their corresponding regions produced in the previous step (Fig. 5(e)). As illustrated in Figure 5(e) some small regions are created. The process is followed by computing new MBR regions and subtracts them from their corresponding regions. This hierarchical procedure is continued until "no regions" are produced any more. That means the progress stops when either no new regions created any more or the size of produced regions is less than predefined threshold. After convergence the final polygon is

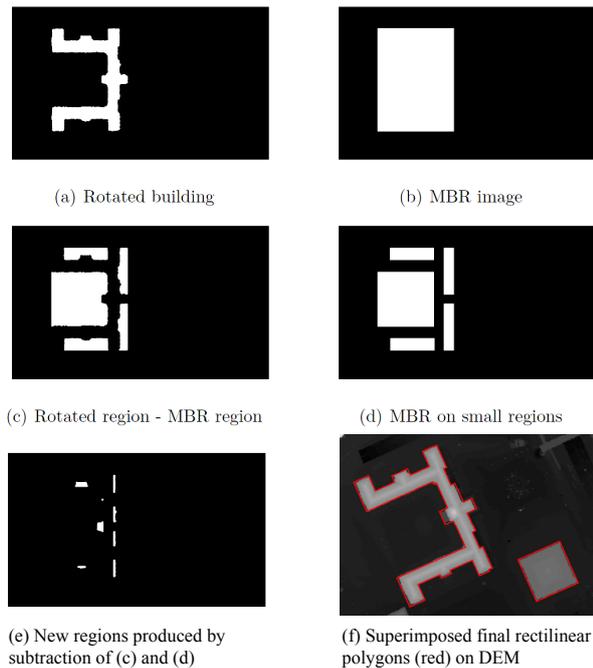


Figure 5: MBR based polygon approximation

rotated back to original orientation (Fig. 5(f)). In this figure the red lines highlight the rectangular polygons.

- If the building is not rectilinear, i.e., at least one side is not perpendicular to the other sides, a method based on Random Sample Consensus – RANSAC (Fischler and Bolles, 1981) is applied for approximation.

RANSAC was originally devised to robustly fit one single model to noisy data. It turns out, however, that it can also successfully be used to fit a beforehand unknown number of models to the data: In the case of the ground plan boundaries the number of line segments is initially unknown. We simply apply the method repeatedly – always deleting the already fitted given points from the input data – until either:

- a) we consider the lines found so far sufficient to construct the ground plan completely or
- b) the number of points fitting to the best line segment with respect to the current iteration step falls below a chosen threshold  $t$ .

In this algorithm the straight lines are repeatedly extracted using RANSAC algorithm and merged to form the final polygon. Figure 6 shows the RANSAC based approximation of the same building represented in Figure 5.

As an alternative to the RANSAC based approximation algorithm, a method similar to MBR-based is proposed for the buildings containing several orientation directions. The method called Combined Minimum Bounding Rectangle (CMBR) based algorithm for hierarchical approximation of non-rectangular polygons. In this method, based on each orientation a MBR (rectangle) polygon is estimated as first approximation level, as shown in Figure 7(a and b).

Intersection of the rectangles corresponding to each orientation produces the first approximation of non-rectangular building (Fig.

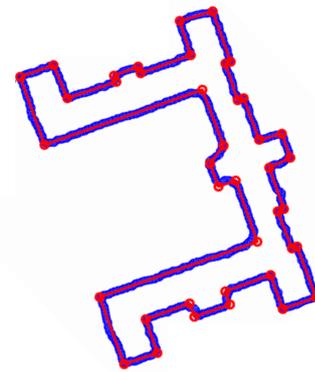


Figure 6: Approximation of polygon obtained using RANSAC

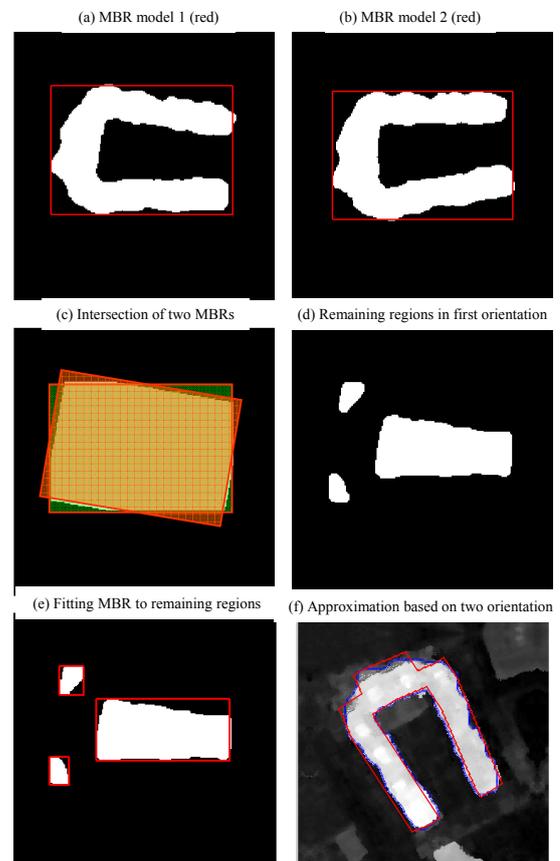


Figure 7: CMBR based polygon approximation

7(e), yellow region). The first approximation area is subtracted from original binary region to generate the remaining regions which should be approximated. The process is continued, similar to the MBR-based method but using all orientation directions, until no more regions remain or remaining regions contain small number of pixels. Figure 7(f) illustrates the final approximation result of the sample building by using two main orientations. The approximation result could be improved by taking into account more significant orientations.

The proposed algorithms for approximation of the building outlines and finally, generating 3D prismatic models have been implemented in a city area containing flat roof buildings in Tunis (Fig. 1). For this experiment, two approaches of MBR- and

CMBR-based, have been utilized to approximate buildings containing rectangular and non-rectangular outlines.

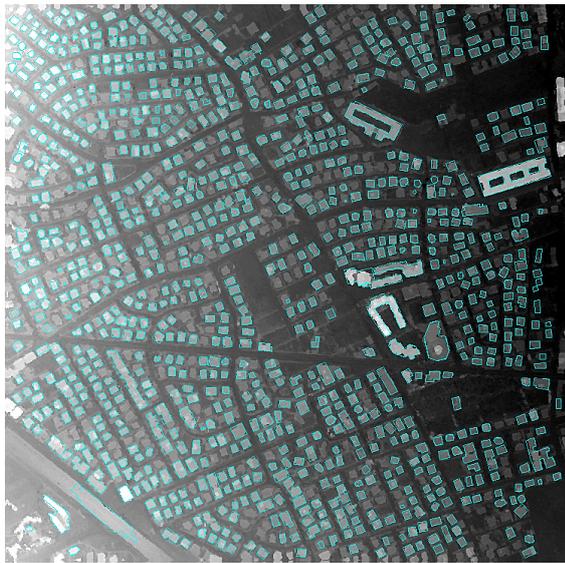


Figure 8: Building outline approximation (green polygons)



Figure 9: Approximation result (Binary regions) (in detail view)

Figures 8 and 9 show the approximation results of the extracted building outlines using classification technique presented in this paper. For a final 3D reconstruction of the buildings, an average height is estimated for each building to shape prismatic model. For each building, the prismatic model is shaped using average height and polygon nodes which are extracted by approximation approaches. The polygons are merged to form prismatic models as illustrated in Figure 10. Additional 3D representation is provided by superimposing the ortho rectified image of corresponding area to the 3D prismatic models (Fig. 11).

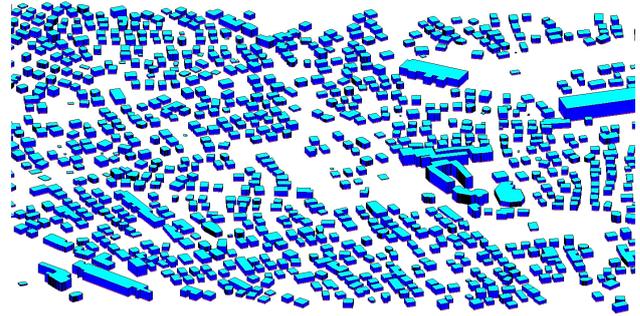


Figure 10: 3D prismatic models

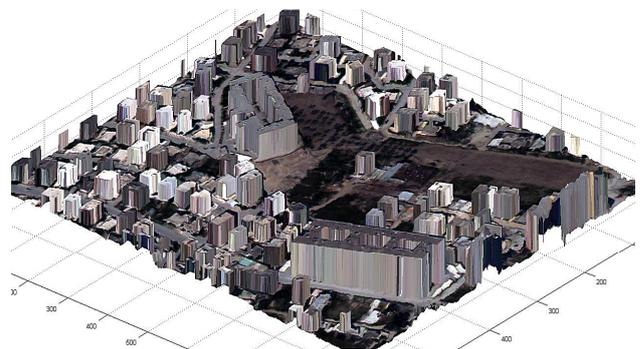


Figure 11: Superimposing of ortho rectified image on 3D prismatic models (zoomed to a smaller region)

#### 4 EXPERIMENTS

Finally, in this section we compare three-dimensional urban region reconstruction performances of the proposed two different approaches. Using two large test area images and DSMs of Tunis city (both in  $2000 \times 2000$  pixel sizes), we have performed our quantitative performance comparisons. For testing shape detection performances, we applied two different analysis; object based shape detection performance and pixel based shape detection performance. The object based shape detection performance considers a building as correctly detected if the algorithm can detect any shape on it regardless of the detected shape. Based on object based performance analysis, active shape model based approach detected 1023 of 1151 buildings correctly (88.87% true positive percentage) and 20 buildings are detected in non-built areas (1.00%). However, the prismatic model based approach detected 953 of the 1151 buildings correctly (82.79% true positive percentage) and 0 buildings are detected in non-built areas. The pixel based performances are computed by considering how many groundtruth building pixels are detected correctly. The active shape model based approach detected 82.12 % of building pixels correctly, and 10.08 % of the pixels in the result shape are falsely labeled as buildings in non-built areas. However, the prismatic model based approach detected 96.26 % of building pixels correctly, and 18.54 % of the pixels in the result shape are falsely labeled as buildings in non-built areas. Shape detection results show that, the active shape model based approach is more successful to detect building structures, and the prismatic model can miss building locations if the building shape cannot be represented with prismatic models. However, if the building location is correctly detected, the prismatic approach can estimate more accurate shape for buildings since it does not contain discontinues as the active shape based approach has in the connected active shape models. For testing height estimation performances,

the best approach is to use LIDAR data of the same region for comparison. Unfortunately, for the test region the LIDAR data does not exist. Therefore, the height estimation of the results are checked by comparing the mean of building height differences from the ground in result data and in nDSM data. In this comparison, the active shape model based approach gave 0.586 meter difference value, and the prismatic model based approach gave 0.724 meter difference value. The low differences of the building height values from nDSM data indicates the reliable results of the proposed automatic approaches in building height assignment steps.

## 5 CONCLUSIONS

Herein, we introduced two different approaches for automatically 3D city model generation using DSMs which are obtained from very high resolution satellite images. Besides proposing new approaches for 3D model generation, we provided quantitative comparisons of the 3D models based on building shape detection and height estimation performances. In order to give an insight view to the reader, we also discussed computation time requirements and implementation difficulties of those approaches. To test our algorithms we used two test areas which have completely different structuring types. We used DSMs obtained over Munich and Tunis cities by using WorldView-2 satellite sensors. However, the final assessment prove that the methodologies lead to very good results. We believe that the results can also assist the applications like detailed city monitoring, change detection, urban structure analysis, planning, damage investigations, and population assessments.

## REFERENCES

- Arefi, H., Engels, J., Hahn, M. and Mayer, H., 2007. Approximation of building boundaries. In: Urban Data Management Systems (UDMS) Workshop, Stuttgart, pp. 25 – 33.
- Arefi, H., Engels, J., Hahn, M. and Mayer, H., 2008. Levels of detail in 3d building reconstruction from lidar data. In Proceedings of International Archives of Photogrammetry, Remote Sensing, and Spatial Information Sciences 37, pp. 485–490.
- Brunn, A. and Weidner, U., 1997. Extracting buildings from digital surface models. In Proceedings of International Archives of Photogrammetry, Remote Sensing, and Spatial Information Sciences.
- Canu, D., Gambotto, J. and Sirat, J., 1996. Reconstruction of building from multiple high resolution images. In Proceedings of International Conference on Image Processing.
- Fischler, M. A. and Bolles, R. C., 1981. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. Communications of the ACM 24(6), pp. 381–395.
- Fradkin, M., Roux, M., Maitre, H. and Leloglu, U., 1999. Surface reconstruction from multiple aerial images in dense urban areas. In Proceedings of IEEE Computer Vision and Pattern Recognition 2, pp. 262–26.
- Harris, C. and Stephens, M., 1988. A combined corner and edge detector. Proceedings of 4th Alvey Vis. Conf. pp. 147–151.
- Jacobsen, M., 2005. Point-process theory and applications: Marked point and piecewise deterministic processes (Probability and its applications). First edition edn, Birkhaeuser Boston.
- Kolbe, T. H., Gröger, G. and Plümer, L., 2005. CityGML - Interoperable Access to 3D City Models. In: International Symposium on Geo-information for Disaster Management, Springer Verlag.
- Maas, H., 1999. Fast determination of parametric house models from dense airborne laserscanner data. ISPRS Workshop on Mobile Mapping Technology, Bangkok, Thailand.
- Ortner, M., Descombes, X. and Zerubia, J., 2002. Building extraction from digital elevation models. INRIA Research Report.
- Sirmacek, B., Taubenboeck, H., Reinartz, P. and Ehlers, M., 2011. Performance evaluation for 3-d city model generation of six different dsms from air- and spaceborne sensors. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 1 (99), pp. 1939–1404.
- Tournaire, O., Bredif, M., Boldo, D. and Durupt, M., 2010. An efficient stochastic approach for building footprint extraction from digital elevation models. ISPRS Journal of Photogrammetry and Remote Sensing 65 (4), pp. 317–327.
- Valero, S., Chanussot, J. and Gueguen, P., 2008. Classification of basic roof types based on vhr optical data and digital elevation model. IEEE International Geoscience and Remote Sensing Symposium IGARSS'08, Boston, USA.