

# SEGMENTATION AND CLASSIFICATION OF POINT CLOUDS FROM DENSE AERIAL IMAGE MATCHING

Mohammad Omidalizarandi<sup>1</sup> and Mohammad Saadatseresht<sup>2</sup>

<sup>1</sup>University of Stuttgart, Germany  
mohammadzarandi@gmail.com

<sup>2</sup>University of Tehran, Iran  
msaadat@ut.ac.ir

## **ABSTRACT**

*In the recent years, 3D city reconstruction is one of the active researches in the field of photogrammetry. The goal of this work is to improve and extend surface growing based segmentation in the X-Y-Z image in the form of 3D structured data with combination of spectral information of RGB and grayscale image to extract building roofs, streets and vegetation. In order to process 3D point clouds, hybrid segmentation is carried out in both object space and image space. Our experiments on three case studies verify that updating plane parameters and robust least squares plane fitting improves the results of building extraction especially in case of low accurate point clouds. In addition, region growing in image space has been derived to the fact that grayscale image is more flexible than RGB image and results in more realistic building roofs.*

## **KEYWORDS**

*Object Surface Segmentation, Image Segmentation, Classification, Region Growing, X-Y-Z Image, Intensity*

## **1. INTRODUCTION**

In the current state-of-the-art, integrating the images with 3D data like range images incorporates computer vision and photogrammetry in this area of research. Automatic 3D reconstruction from 3D point clouds or range image is still one of the active research areas that has many applications in forestry, urban planning, tourist information systems and so on.

3D data measurements are grouped with respect to their similarity measures to define meaningful, coherent and connected segments. Acquisition of the high accurate and dense 3D data can assist us in the direction of automatic extraction of building models. Image matching can be considered a renaissance in the modern photogrammetry due to generating dense and high accurate point clouds with low price in comparison with LiDAR data. Recently, many algorithms are focused on the extraction of planar surfaces especially extraction of roof facets for 3D building reconstruction that related works will be reviewed in the next section. However, building extraction is still challenging issue due to the complex building roofs, occlusions and shadows. Availability of the 3D structured data for each pixel and intensity values from the high resolution aerial images in addition to usage of high performance computer can assist us to deal with the challenging issues and achieving better segmentation results with low price. Thus, this combination can be considered the important step to reach the goal of automatic 3D reconstruction and object recognition. The goal of this work is to improve and extend surface growing based segmentation in the X-Y-Z image in the form of 3D structured data with combination of spectral information of RGB and grayscale image to extract building roofs, streets and vegetation.

Jarżabek [1] proposed segmentation algorithm based on updating plane parameters and merging surface patches which works well in case of high accurate point clouds. Therefore, we improved the aforementioned algorithm by considering robust estimation and removing outliers. In addition, no need of merging surface patches due to considering unlimited nearest neighboring points. However, it is quite time consuming and can be used merely for building roofs extraction.

Rabbani et al. [2] proposed segmentation of point clouds using smoothness constraint. This approach worked well for the high accurate point clouds. However, it seems problematic in case of low accurate point clouds. Therefore, in this work, it has been implemented to segment streets to increase the run time of the implemented program and preventing updating the plane parameters.

Building roofs extraction is the most important feature in our segmentation procedure. But, in some cases, there is an overlap between the building roofs and trees. Thus, it enforces us to classify trees and building roofs to achieve better result of building roofs extraction. In this work, vegetation is extracted by means of spectral information from RGB image. Generally speaking, in RGB image, vegetation is illustrated by the green color. However, this green color is not the same in all pixels of the RGB image. Therefore, we consider the threshold for the green channel of the RGB image and considering those pixels as vegetation which the differences of two other channels from green channel are above the pre-defined threshold. Filtering the extracted vegetation eliminate noisy pixels from the extracted vegetation and fill the gaps between them. In this work, we applied median filter with mask size 14 to remove the noises from the image and achieve better result of building and street extraction. Finally, morphologic dilation can be utilized to fill the gaps between extracted vegetation.

## **2. RELATED WORKS**

Segmentation as an important part of data processing is applied on the various kinds of data sets like 3D point clouds or range image to partition data sets into meaningful, disjoint and connected segments with homogenous property. Hence, the purpose of the segmentation is to group points or pixels with similar features into segments. Segments are smooth surfaces that are achieved by grouping neighboring points or pixels with similarity measures, such as the direction of a locally estimated surface normal or intensity values of each pixel. Generally speaking, point cloud segmentation can be considered a difficult subject, especially in presence of the noise. In addition the gaps between point clouds (mismatch points from image matching for each image pixel) and varying point densities make it more problematic.

Methods of surface extraction can be categorized in two main groups. Firstly, surface parameters can be estimated directly by clustering or finding maximum parameter in the parameter space. Secondly, point clouds can be segmented on the basis of proximity of the point clouds or similarity measures like locally estimated surface normals [3]. In other word, range segmentation problem can be categorized in two main approaches: region-based and edge-based segmentation problem. Furthermore, region-based segmentation problem can be divided in two main groups: parametric model-based segmentation algorithms and region-growing algorithms [4].

### **2.1. Edge-Based Segmentation**

Edge based segmentation is a special case of region growing algorithms since points are bounded within the closed boundaries and connectivity between them is not on the basis of spatial relations [5]. The algorithm starts with extracting the edges along the boundaries of different regions. Edges can be detected while changes in local surface properties (surface normals, gradients,

principal curvatures) exceed a pre-defined threshold. This procedure followed by grouping the points inside the boundaries and result the segmented regions.

As a drawback of edge-based segmentation algorithm, in many cases, they generate non-closed boundaries. Moreover, it is difficult to detect discontinuity in the curved surfaces due to smoothness in this kind of surfaces which lead to under-segmentation in the range image [4]. Furthermore, only measurements close to the edges are taking into account and other available measurements are not considered as well [6]. In addition, these algorithms are very sensitive to the noise in the range data.

## **2.2. Clustering**

Clustering represents another approach of range image segmentation in the feature space and can be considered one of the main categories of the region-based segmentation problem. In clustering approach, parameters of the surfaces are introduced and then point clouds are grouped based on surface parameters. This algorithm is performed by subdividing the points into disjoint regions with homogeneous property. The points inside each region share similar property which is different from other regions and therefore distinguish each region from other regions.

Filin [7], Hoffman and Jain [8] and Filin and Pfeifer [9] describe different approaches of clustering.

Clustering is similar to Hough transform approach in case of working in the feature space. However, unlike the Hough transform, the attributes are calculated locally and the risk of grouping points that are not connected is very low [7]. Moreover, unlike the Hough transform, in clustering approach, with consideration of proximity, just points with similar surface normals are clustered together [9].

## **2.3. 3D Hough Transform**

3D Hough transform can be considered the subset of clustering approach. The purpose of the Hough transform technique is to find objects by a voting procedure. This voting procedure is carried out in a parameter space (Hough space). In 3D Hough transform, high accumulator value gives the hypotheses for detecting planes.

As a drawback of this approach, points' connectivity is not taken into the account and many spurious plane surfaces may be extracted from those points that are not in the same plane with the given point [9]. Moreover, precision or resolution of parameter space seems problematic in this approach.

## **2.4. Geometric Primitives-Based Segmentation**

Geometric primitives-based segmentation belongs to clustering approach. This method is suited for fitting higher-order surfaces to the range data and not just planar regions. Bab-Hadiashar and Gheissari [4] describe parametric surface model-based segmentation from the range image based on Surface Selection Criterion (SSC) that is carried out by minimizing the strain energy of the thin surface. In order to choose the appropriate surface model, surface is fitted to the measurements and sum of squared residuals between the fitted surface and range data will be small. The perfect fitness of the surface to the measurements is achieved by bending and twisting the surface till it will be closer to the measurements.

As an advantage of the parameter model-based approaches is that the segmentation and object recognition is carried out at the same time [10].

As a drawback of the geometric primitives-based segmentation is that the number of surface types directly affects on the segmentation results [11].

## 2.5. Surface Growing-Based Segmentation

Surface growing in object space is the equivalent of region growing in image space. Surface growing in object space can be performed by grouping point clouds which are spatially close and share similar measure properties like the direction of a locally estimated surface normal, gradient and the principal curvatures. As a result of surface growing method, point clouds are segmented into multiple surfaces. Surface growing algorithm starts from the optimal seed points and surfaces extend to neighboring point clouds based on pre-defined criteria. In this fashion, selection of the seed point is important step and final segmentation results are dependent on it. In order to find optimal seed point, firstly, plane equation is defined for each seed point and its neighboring point clouds. Then, the residuals (orthogonal distances of the point clouds to the best fitted plane) are computed. The point within the fitted plane with lowest square sum of residuals is considered the optimal seed point. Outliers would affect on the results of surfaces normal and consequently on the square sum of residuals which leads to a failure of detecting proper seed points of surfaces. Therefore, robust least squares adjustment is applied to detect the optimal seed point even in the presence of outliers. In this approach, firstly, the plane fit to the surface points and their neighborhood. The candidate point that its orthogonal distance to the fitted plane is below the pre-defined threshold is accepted as a new surface point. Due to increase the efficiency of the program while using low accurate point clouds, the plane equation is updated after adding the new candidate point to the corresponding surface points. The neighborhood threshold and residual threshold are used to determine the smoothness of the fitted plane. Secondly, the local surface normal at each point is compared with its neighboring points. The neighboring points are accepted if the angle between its surface normal and normal at the neighboring point is below the pre-defined threshold.

As a privilege of surface based algorithms in comparison with edge based algorithm is no need to identify the surface boundaries at the preliminary step. Due to their easy implementation and well time performance, surface growing algorithms are the prevailing method for point cloud segmentation. However, the weakness of this approach is connected with a proper choice of the seed points. Furthermore, selecting different seed points may results different segmentation regions [12]. In addition, this algorithm tends to generate distorted boundaries due to segment objects in the region level instead of pixel level [4].

## 2.6. Scan Line-Based Segmentation

Scan line-based segmentation is the subset of region growing approach. It can be implemented in two following stages: [3]

- (a) Splitting scan lines into small pieces
- (b) Merging scan lines with neighboring scan lines on the basis of similarity conditions.

Jiang and Bunke [13], Sithole and Vosselman [14] and Han et al. [15] describe different approaches of scan line-based segmentation.

As drawback of general scan line segmentation approaches, point clouds need to read in parallel scan lines due to increasing the efficiency of the algorithms. In addition, this approach tailored to

extract planar surfaces in 2.5D range image and there is no straight forward solution for 3D point clouds [2].

## **2.7. Curvature-Based Segmentation**

Curvature-based segmentation belongs to region growing approach. Surface curvature like surface normal as local surface properties can be utilized in object recognition. Surface curvature can be considered the rate of surface deviation from its tangent plane and can be estimated from local neighborhood on the surface. Furthermore, it does not change due to altering viewing direction and occlusion [16]. Hoover et al. [17] describes curvature based segmentation based on Gaussian and mean curvatures. In this method, Gaussian (H) and Mean (K) curvature are computed for each pixel in a certain mask size. Then, pixels are labeled with combination of sign values of H and K. Curvature values are classified to negative, zero and positive values. Afterwards, neighboring pixels with similar labels are grouped from seed regions.

As a drawback of this method is that in case of noisy data, it leads to over-segmentation because curvature values can be not computed accurately. In addition, in the presence of outliers in the region growing procedure, applying just continuous curvature values as a condition leads to merge separate regions [11]. Another disadvantage of this method is that higher order surfaces are limited to the pre-defined number of different surfaces and other types of surfaces are not recognizable [4].

## **2.8. Merging Triangulated Irregular Network Meshes**

Triangulated Irregular Network (TIN) meshes algorithms belongs to region growing approach. Gorte [18] describes a variation that was inspired by a raster-based algorithm but carried out on triangulated laser point clouds. This Method uses TIN meshes to reconstruct the surface.

As a privilege of this fashion, working directly on triangulated laser point clouds instead of rasterization and avoiding discretization of the data sets and loss of information.

## **3. THE PROPOSED SURFACE GROWING-BASED SEGMENTATION**

In this work, surface growing based segmentation approach is implemented on the basis of bottom-up strategy. It means that procedure starts from the some selective seed points and the segments are grown based on pre-defined similarity criterion. The selection of the seed points is one of the important issues in this process that can affect directly on the final results. Segmentation is carried out in both object space and image space to extract building roofs and streets. Firstly, it is applied in the object space based on surface growing-based segmentation by using surface normals that are computed from X-Y-Z image for each valid image pixel. In object space, adjacent point clouds are grouped based on similarity measures like direction of locally estimated surface normals and distances of the point clouds to the best fitted plane. Afterwards, segmentation proceeds in image space by the usage of segmented regions from the object space and grouping adjacent image pixels based on similarity of gray or color value of neighboring image pixels. Therefore, final segmentation results in extraction of building roofs and streets that extracted regions significantly differ from each other due to having different surface slopes or being in different elevations.

The segmentation method has two steps, Surface normal estimation and region growing that is described and discussed in next sections.

### 3.1. Data Structure

The 3D data sets of this work were provided from pixel-wise image matching. Firstly, parallax is generated as a result of image matching and consequently 3D coordinates are produced for each image pixel. Parallax is represented by so-called disparity image and 3D coordinate is stored in X-Y-Z image. Thus, corresponding X-Y-Z image is adapted to range image segmentation problem. In previous methods of segmentation from LIDAR data, 3D point clouds are transformed into the raster domain to be adapted for range data segmentation problem that causes loss of information. However, generated X-Y-Z image provides us 3D point cloud for each pixel that is 3D structured data and no need of the transformation of 3D data from 3D space to  $2\frac{1}{2}D$  space. Moreover, additional information like gray or color value from gray image or RGB image can be utilized to interpret each pixel as well. Furthermore, valid or invalid image pixels are the result of matching process and 3D coordinates are merely available for valid image pixels. Hence, validity of 3D coordinate for each X-Y-Z image pixel can be checked by another image that is so-called valid image. In addition, DTM can be generated by filtering the DSM and it is used to distinguish ground or non-ground range data.

### 3.2. Surface Normal Estimation Using Nearest Neighboring Search Algorithm with Robust Estimation

Jarżabek [1] computed surface normals in the object space by using LiDAR data. However, we extended the algorithm to the X-Y-Z image pixel and improved it by considering the robust estimation to mitigate the effects of outliers.

The surface normal for each X-Y-Z image pixel is estimated by fitting a plane to neighboring X-Y-Z image pixel which contains 3D coordinates for each image pixel. In this method, nearest neighboring search algorithm is applied to find closest points to the query point. The fitted plane is affected by the noises especially when outliers exist in the data sets or selected 3D point clouds are not in the same plane. Using robust estimation assists us to discard outliers and recalculate the surface normal again. In this method, mask size 5 is defined as a default mask size and nearest neighboring search starts. If it can not find pre-defined number of neighboring X-Y-Z image pixel due to invalidity of image pixels, the mask size will be resized to the next larger size and nearest search algorithm repeat. This process proceeds to reach pre-defined number of neighboring X-Y-Z image pixel. The process of calculating surface normals is carried out with the usage of nearest neighboring search algorithm to define neighborhood and using robust estimation to discard outliers and recalculating the surface normals again.

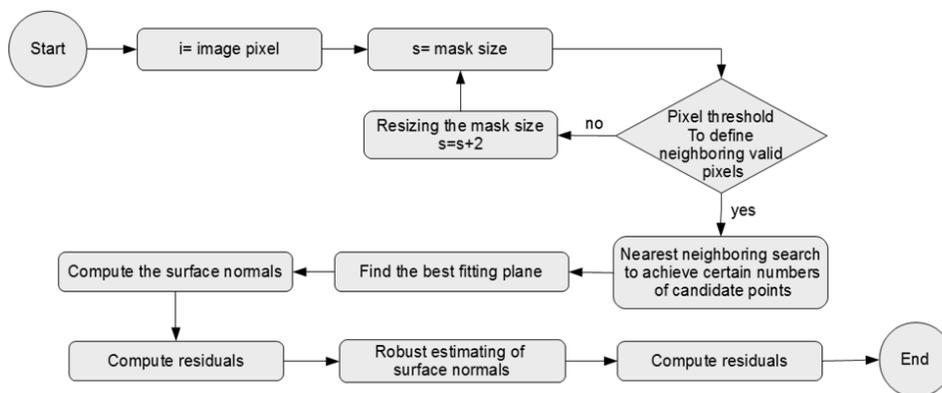


Figure 1. Flowchart of robust surface normal estimation

### 3.3. Nearest Neighboring Search Algorithm

Nearest neighboring search problem, also known as proximity search problem, is the procedure of finding closet points to the query point that is carried out in metric space. In this work, we determined the nearest neighboring search algorithm on the basis of 3D Euclidian distances of the query point to its neighbors. This algorithm starts by defining the minimum number of nearest neighboring points as a threshold. The trivial way is to compute distance of the query point to the whole data sets but, it is time consuming. Thus, in order to increase the efficiency and run time of the algorithm, firstly, mask size 5 is defined and numbers of valid point inside the pre-defined mask size is counted. If the number of counted point would be below the threshold, the mask size is resized to mask size 7. This procedure proceeds until the number of points within the mask size be equal or even more than pre-defined threshold of the nearest neighboring points. Afterwards, the position of query point in the object space and obtained mask size from the previous step is utilized in nearest neighboring search algorithm. Next, the 3D Euclidian distance of the query points to the candidate points within the mask size is calculated. Then, candidate points are sorted in the descending order based on their 3D Euclidian distance values to the query point and stored in the vector. Thereafter, pre-defined numbers of neighboring points are picked from the last components of this vector and enter to the surface normal function to calculate surface normal based on nearest neighbors.

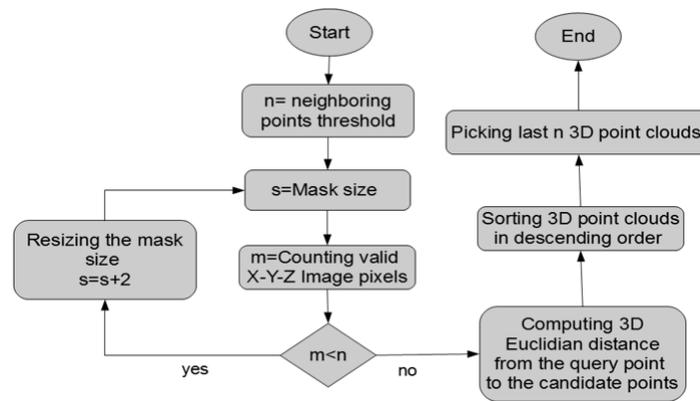


Figure 2. Flowchart of nearest neighboring search algorithm

### 3.4. Robust Plane Fitting

Plane fitting is the crucial process for extracting plane features from 3D point clouds. In order to achieve optimal plane parameters, Least Square Fitting (LSF) is applied. Furthermore, Robust Estimation has been applied to reduce the effect of noises.

The least squares fitting approach can result to optimal plane fitting parameters while there are no outliers in the datasets. This approach may lead to failure in case of noisy points or isolated points that are far from the plane. Based on the residuals (orthogonal distances of 3D point clouds to the fitted plane) from the previous adjustment, Danish robust estimation gives each observation a suitable weight using equation (1). In this work, at the first, surface normal is calculated using least squares fitting (LSF). Then orthogonal distances of 3D point clouds to the fitted plane are calculated. Note that normal vectors are already normalized and norm of them are equal to 1. In the next step, standard deviation of calculated orthogonal distances from previous step is calculated. If an orthogonal distance is within 2 times of the calculated standard deviation, the point will be given a weight one. Otherwise, it will be given a weight zero to discard from our

calculation. In the next step, weights are multiplied by 3D point clouds and surfaces normal are recalculated again to get better results.

$$w_i = \begin{cases} 1 & |d_i| \leq 2\sigma \\ 0 & |d_i| > 2\sigma \end{cases} \quad (1)$$

$$E[|d_i|] = 0 \quad (2)$$

Where  $w_i$  is weight of each 3D point cloud,  $d_i$  is an orthogonal distance of 3D point clouds to the fitted plane and  $\sigma$  is standard deviation.

#### 4. REGION GROWING-BASED SEGMENTATION

In this work, segmentation method is applied based on region growing. This approach can be applied either in object space or in image space. In object space, grouping points is carried out on the basis of geometrical relations of neighborhood like surfaces normal differences and in image space, grouping is performed on the basis of spectral relations of neighborhood like grey value differences. The first step in region growing is to select a set of seed points. The initial region starts from optimal seed points. Then, regions are grown from these seed points to neighboring points depending on a similarities criterion. The criterion in object space can be the orthogonal distances of the 3D point clouds from the plane and angle threshold between surface normal of the seed point and surface normals of the neighborhood. In addition, the criterion in image space can be pixel intensity.

##### 4.1. Algorithms of Region Growing in Object Space and Image Space

The algorithms use the adjacency relations in the X-Y-Z image where X, Y and Z values are available for each pixel. In order to implement region growing in object space, we use point normals and their residuals which already assigned to each image pixel. In addition, implementing region growing in image space can be carried out by using intensity values of image pixels. This algorithm in object space aims to group points which locally connected and belong to smooth surface. Furthermore, in image space, it aims to group image pixels with nearly similar intensity values. The points in a segment should locally make a smooth surface, whose normals do not vary too much from each other. In this procedure, angle threshold ( $\theta_{th}$ ) between the selected seed point and its neighboring points is applied to add qualified points to the region. In the next step, using residual threshold ( $r_{th}$ ) can lead to detect smooth surfaces in different height level [6].

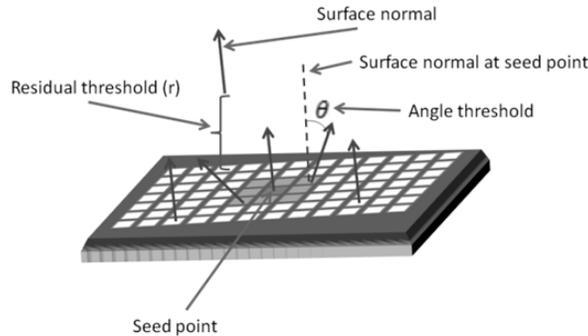


Figure 3. Representation of surface normals and segmentation parameters in object space

In this work, in order to calculate surface normals using robust estimation, the segmentation has been implemented in two different methods.

First method can be used to extract objects (building roofs and streets) for dense and high accurate point clouds. In this method, selected seed point is compared with its neighboring points with considering angle and residual threshold. In the next step, one of neighboring point can be considered as a seed point and this procedure is repeating. First approach leads to quite good results due to use dense and high accurate point clouds. This method, even can work to extract streets for low accurate point clouds because the surface normals in the flat areas do not vary too much. Extracting street by using this method can improve the run time of the program.

Second method can be used to extract building roofs for low accurate point clouds. In this method, optimal seed point is selected and it is compared with its neighboring points with considering angle and residual threshold. In the next step, those neighboring points that are qualified are added to surface points. Then, plane is updated and new plane parameters are calculated. Updating the plane parameters, can increase the efficiency of the algorithm in comparison with the first method. In this approach, we need to consider more neighboring points in comparison with the first approach due to use low accurate point clouds. Consequently, updating the plane parameter and considering more neighboring points can increase the run time of the program.

These approaches are investigated in different case studies by using RGB image, grayscale image and both of them. Using RGB image leads to extract building roofs with more detail but, the shape of extracted building roofs are not close to their realistic shape. On the other hand, using grayscale image results in more realistic building roofs and approximately can cover the whole building roofs due to more flexibility of gray values in region growing procedure. Segmentation in RGB image can be carried out by comparing three channels of the pixels in the image. Furthermore, it can assist us to detect vegetation from images based on comparing three channels and considering those pixels as vegetation which the differences of two other channels from green channel is above the pre-defined threshold.

DTM is also used as additional information to recognize ground and non-ground 3D point clouds. It can be applied by considering the nDSM threshold and separate point clouds to two main categories of terrain surfaces and building roofs.

In the first method of calculating surface normals using robust estimation, dense and high accurate point clouds in company of grayscale image is applied and segmentation algorithm based on region growing includes following steps:

- (1) Sorting point clouds with respect to their standard deviation of residuals to the fitted plane in descending order. (Note: each image pixel contains 3D coordinates in object space, 2D coordinate in image space, point normal and its standard deviation of residual).
- (2) Picking the seed point with minimum standard deviation of residual and remove it from seed points.
- (3) Defining the mask size and nearest neighbors threshold.
- (4) Performing nearest neighboring search algorithm to find valid neighbors with minimum Euclidian distances.
- (5) Comparing of numbers of valid neighbors with nearest neighbor threshold. If it reaches to defined numbers of neighbors, goes to next step. Otherwise, resizing the mask size and repeating the process from step 4.
- (6) Computation of angle between normal vector of selected seed point and normal vectors of other neighboring image pixels.

- (7) Picking neighboring image pixels with values less than angle threshold ( $\theta_{th}$ ) and residuals threshold ( $r_{th}$ ).
- (8) Considering all similar neighboring image pixels as one segment and coloring them in another image. (Note: all similar neighboring image pixels store in two real and virtual segments. Real segment to keep all image pixels with similar properties and virtual segment to pick the seed point and then remove it from that segment)
- (9) Picking the last component of the created segment as a new seed point and remove it from virtual segment.
- (10) Repeating the process from step 4 and add neighboring image pixels to the created real segment. Keeping on this step until virtual segment becomes empty and no more seed point can be added to the created real segment. (Note: Before choosing the seed point, checking the intensity value of the colored image and make sure that selected seed point and its neighbor have not selected yet).
- (11) Adding created segment to the region list and keeping on this process from step 2 until no more image pixels remain in the colored image without coloring and segmentation.
- (12) Coloring segmented regions in another image. In this approach, in order to visualize colored segmented regions, considering the segment size threshold to avoid coloring small segment regions.



Figure 4. Representation of seed point and its neighboring points in first method

In second method of calculating surface normals using robust estimation, three types of point clouds with different level of density and accuracy in company of RGB, grayscale image and DTM is utilized and segmentation algorithm based on region growing includes following steps:

- (1) Sorting point clouds with respect to their standard deviation of residuals to the fitted plane in descending order. (Note: each image pixel contains 3D coordinates in object space, 2D coordinate in image space, point normal and its standard deviation of residual).
- (2) Picking the seed point with minimum standard deviation of residual and remove it from seed points.
- (3) Considering nDSM threshold to categorize current seed point to extract building roofs or to extract streets.
- (4) In case of building roof extraction, current algorithm is performed. Otherwise, in case of street extraction, first algorithm is applied.
- (5) Defining the mask size and nearest neighbors threshold.
- (6) Performing nearest neighboring search algorithm to find valid neighbors with minimum Euclidian distances.
- (7) Comparing of numbers of valid neighbors with nearest neighbor threshold. If it reaches to defined numbers of neighbors, goes to next step. Otherwise, resizing the mask size and repeating the process from step 6.
- (8) Computation of angle between normal vector of selected seed point and normal vectors of other neighboring image pixels.

- (9) Picking neighboring image pixels with values less than angle threshold ( $\theta_h$ ) and residuals threshold ( $r_h$ ).
- (10) Considering all similar neighboring image pixels as one segment and coloring them in another image.
- (11) Updating plane parameters using robust estimation of surface normals with respect to all points that are available in the created segment.
- (12) Considering all neighboring points of all points that belong to the created segment and repeating the process from step 8.
- (13) The updating plane parameter proceeds until no more neighboring points can be added to the created segment.
- (14) Adding created segment to the region list and keeping on this process from step 2 until no more image pixels remain in the colored image without coloring and segmentation.
- (15) Coloring segmented regions in another image. In this approach, in order to visualize colored segmented regions, considering the segment size threshold to avoid coloring small segment regions.



Figure 5. Representation of seed point and its neighboring points in second method

#### 4.2. RGB Image vs. Grayscale Image in Image Segmentation Procedure

RGB image consists of three channels (Red Green Blue) that enable us to extract objects from the images with their real color in the nature like vegetation. On the other hand, grayscale image consist of one channel that is generated from the combination of the three RGB channels. As mentioned before, the objective of this work is to extract building roofs, streets and vegetation. In order to extract vegetation in this work, using RGB image is compulsory. Furthermore, in extraction of building roofs using grayscale image, region segmentation can grow more flexible comparing to the RGB image. For instance, in the table 1, RGB values from the RGB image compares to gray values from the grayscale image.

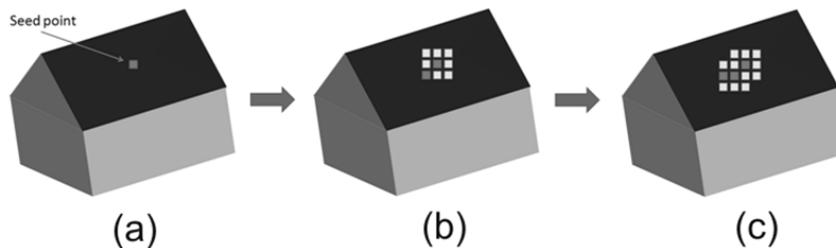


Figure 6. Region growing in image space

Table 1. Example of RGB values vs. gray values in the image segmentation procedure

	Blue	Green	Red	Gray Value
(a)	94	97	88	94
(b)	95	97	89	94
(c)	97	96	90	94

As can be seen from the Table 1, combination of different RGB values can generate the same gray value that demonstrates the higher flexibility of the grayscale image in region growing specially, in extraction of the building roofs. This flexibility leads to better segmentation of the building roofs in the image space.

## 5. BUILDING ROOFS BOUNDARIES

Surface boundaries represent the general shape of the homogeneous regions that is derived by grouping point clouds on the basis of their similarity measures. Building extraction is the main objective of this work. In order to reach this goal, the final result of segmentation in the image space is utilized to delineate the building roofs boundaries. Result of boundary detection in image space from aerial images is better and more accurate than object space from the raw 3D data. It stems from the fact that point clouds do not cover the whole area uniformly and it is noisy, particularly in case of using low accurate point clouds. In this work, two alternative methods are utilized to achieve this goal:

- (1) 2D  $\alpha$ -shape
- (2) Convex hull

2D  $\alpha$ -shape is the generalization of the convex hull of the point sets in the plane. It means while the value of  $\alpha$  reaches to infinity, 2D  $\alpha$ -shapes are equivalent to the convex hull [19]. Furthermore, it can be considered the sub-graphs of the Delaunay triangulation [20]. Furthermore, Parameter  $\alpha$  is utilized to control the level of details of the constructed shapes [21]. The value of  $\alpha$  corresponds to the squared radius of the circles that are fitted to the boundary points.

Convex hull is the minimal convexity of the set of points in the 2D plane and forms the closed polygon. It is typically represented by the sequence of the boundary points in the form of boundary polygon with the minimum convexity.

In this work, in order to apply 2D  $\alpha$ -shapes, CGAL library has been installed and utilized. By the usage of this library, we could detect boundary points that distributed randomly around the extracted building roofs. Therefore, we determined the algorithm to connect randomly distributed 2D points with polyline. Convex hull has been applied by the usage of OpenCV library and no need of additional classes to perform it.

## 6. EXPERIMENTS AND RESULTS

In this work, different types of point clouds with different level of density and accuracy are investigated. Therefore, we carried out our experiments in three case studies.

In the first case Study, we dealt with the dense and high accurate point clouds in addition to grayscale image. In this case study, first method of computing surface normals using robust estimation has been applied. This method leads to reasonable results with dense and high accurate point clouds for the extraction of building roofs. Furthermore, it also works well in the flat areas like streets even with lower accuracy due to not varying the slope of surfaces. During the process of segmentation, there is trade-off between the parameters such as angle threshold( $\theta$ ), residual threshold( $r$ ) and numbers of nearest neighboring point threshold. Furthermore, segment size threshold is considered to remove small segmented regions. These results indicate that inappropriate threshold can lead to over or under-segmentation.

In the second case study, three types of point clouds with different level of accuracy and density in company of RGB image and DTM have been utilized. These three types of datasets are as follows:

- (1) Sparse and high accurate point clouds
- (2) Medium and less accurate point clouds
- (3) Dense and low accurate point clouds

First type of these data sets could not be processed individually because of lack of enough 3D point clouds to define surface normals. However, medium and dense 3D point clouds are processed separately. Furthermore, in order to improve the results of segmentation, these three types of data sets are combined together and results in generating forth type of data sets that is so called hybrid point clouds. In this case study, segmentation procedure like as case study I is performed in object space and image space. However, computation of surface normals is modified by updating planar surfaces that is grown from optimal seed points. As a result, this method leads to come up better extraction of the building roofs with improving the results of surface normals. During the segmentation process in image space, RGB image assists us to extract vegetation on the basis of comparing color values of the image pixels. Moreover, it is also utilized in the process of region growing in the image space. Furthermore, DTM can also be used to recognize ground and non-ground 3D point clouds.

In the third case study, these three types of data sets in addition to RGB image, grayscale image and DTM have been utilized.

In this work, 3D point clouds were available in two file formats: LAS and txt file format. LAS format is the popular format to read and write point clouds. This file format can be read in C++ by the usage of LASreader class that belongs to LAsTools. OpenCV is the proper library to work with the images and matrices with powerful functional operation. Furthermore, it has been used to solve eigenvalue problem and delineation of building roofs boundaries using convex hull. In addition, CGAL library provides us this capability to delineate building roofs boundaries using 2D  $\alpha$ -shape. Segmented 3D point clouds are stored in VRML format to display them in 3D and enables users to walk or fly through the segmented regions.

Figure 7 depicts the results of segmentation in image space without considering the spatial information. Initiative seed points are selected from optimal seed points that achieved from square sum of residuals of point clouds to the best fitted plan in object space. However, regions are grown based on intensity values of image pixels. It means that these results are come out just based on spectral information of the images and no spatial information like locally estimated surface normals are considered. The aim of this experiment is to indicate that region growing based on spectral information without considering the spatial information can not lead to the desirable segmented regions and mostly results in over-segmentation.

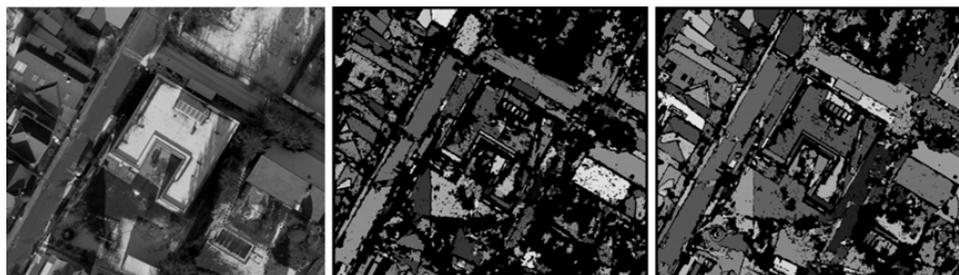


Figure 7. Base image (left), Segmentation in image space without considering the spatial information (middle) (Intensity value threshold<2), (right) (Intensity value threshold<3)

### 6.1. Case Study I

Figure 8 depicts the result of segmentation of surface normals in object space from 3D point clouds in the left side. Moreover, segmentation results in image space by the usage of segmented regions from object space and comparing intensity values of image pixels are indicated in the right side.



Figure 8. Base image (left), Segmentation of surface normals in object space using first algorithm (middle) and image space (right),  $\theta < 5^\circ$ ,  $r < 0.1$  m, 50 nearest neighbours

### 6.2. Case Study II

Figure 9 depicts the result of segmentation in both object space and image space. As can be seen, in this case study, by the usage of RGB image, vegetation has been extracted as additional information comparing to previous case study. Moreover, building roofs has been extracted with more level of details due to comparing color values in three channels of RGB image.



Figure 9. Base image (left), Segmentation of surface normals in object space using second algorithm (middle) and image space (right),  $\theta < 11.5^\circ$ ,  $r < 0.4$  m, 250 nearest neighbors, nDSM threshold = 1.5 [m]

### 6.3. Case Study III

In this case study, RGB image is merely utilized to extract vegetation and grayscale image is used to extract building roofs and streets. Consequently, it stems to the fact that grayscale image is more flexible than RGB image in region growing in image space.

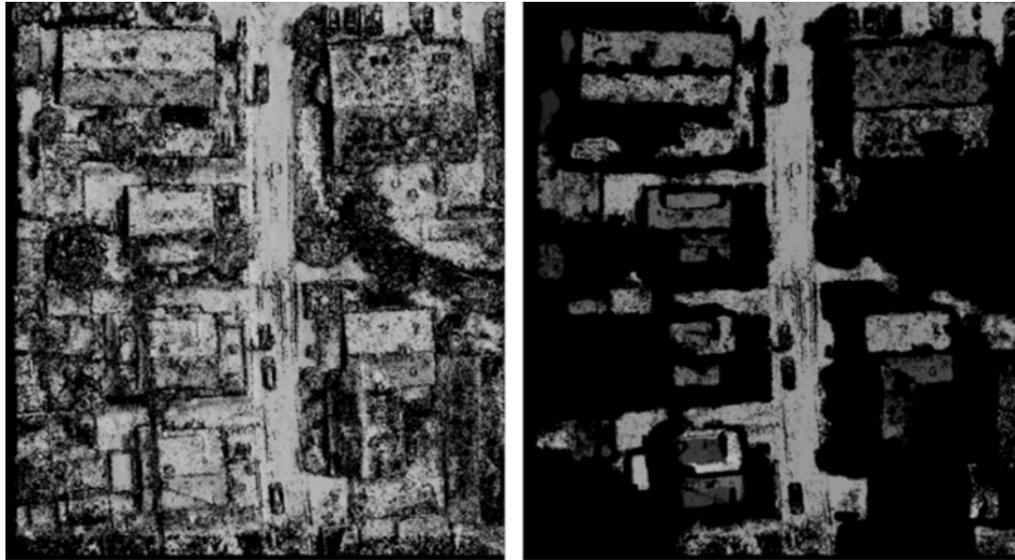


Figure 10. Surface normals (left), Segmentation of surface normals in object space using second algorithm (right),  $\theta = 11.5^\circ$ ,  $r=0.5$  m, nDSM threshold = 1.5 [m]

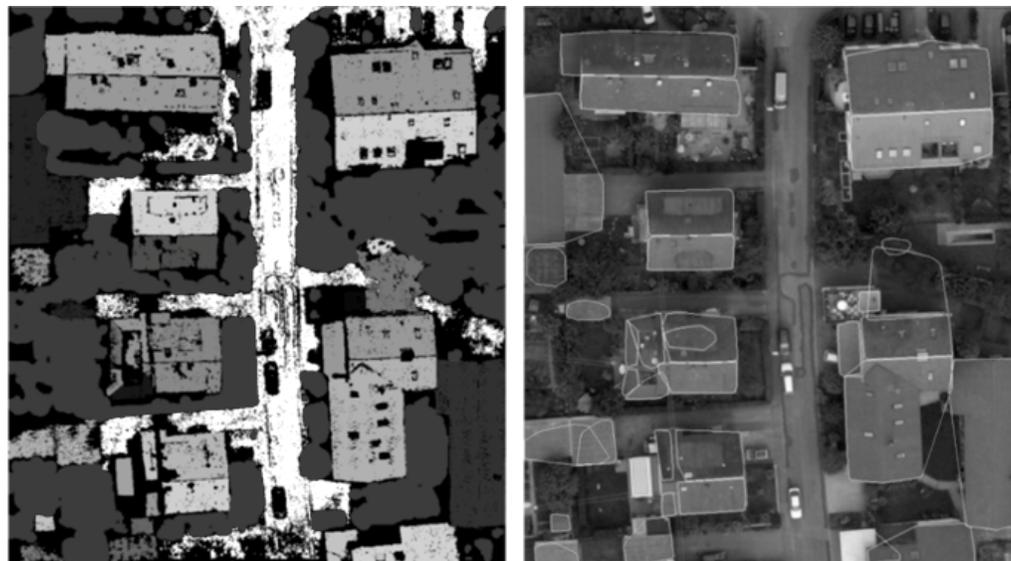


Figure 11. Segmentation in image space (left), Surface boundaries using convex hull (right)

Table 2 represents the evaluation of the validated 3D point clouds in object space. It is carried out by computation of the standard deviation and mean Z shift with respect to the orthogonal distance between the 3D point clouds and the best fitted plane in extracted building roofs and flat terrain.

Table 2. Validation of hybrid point clouds

Surface type	Building roofs	Flat terrain
Standard deviation of validated points	0.409 m	0.473 m
Mean Z shift of validated points	0.0	0.0

Figure 12 depicts the delineation of building roofs using 2D alpha shapes. As can be seen, building roofs boundaries has been extracted more accurately comparing to the convex hull.



Figure 12. Representation of surface boundaries with points (left), with polyline (right) using 2D alpha shapes

Ultimately, 3D visualization of segmented 3D point clouds in VRML format is depicted in Figure 13. The segmented 3D point clouds in each building roofs have been colored with the different color that distinguishes them from other building roofs. Furthermore, vegetation and streets have been colored with the green and gray color respectively.

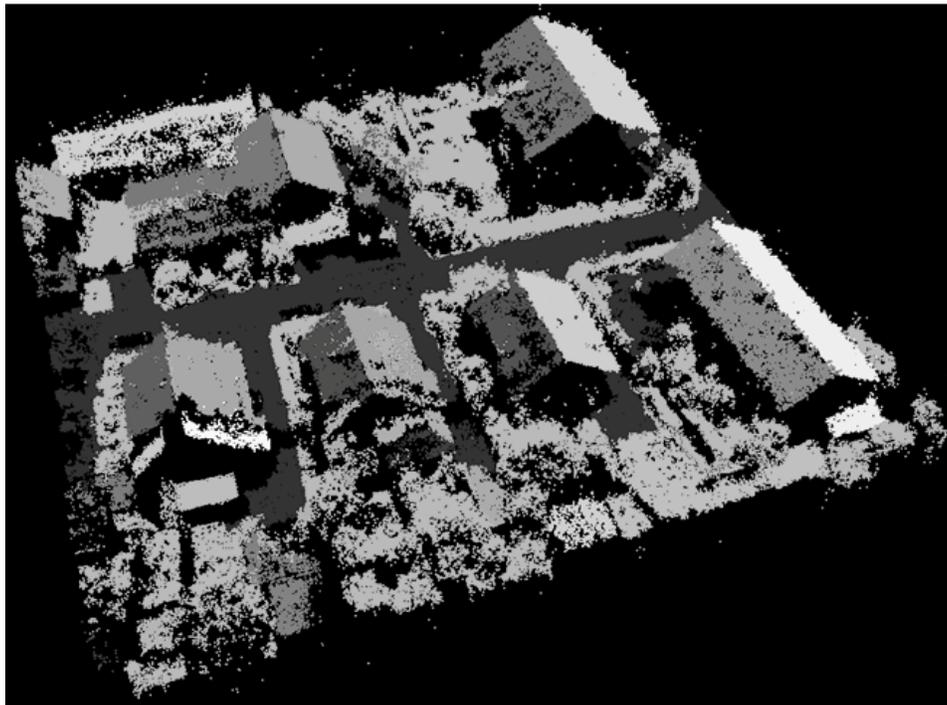


Figure 13. 3D visualization of segmented 3D point clouds in VRML format

## 7. CONCLUSIONS

Robust estimation has been improved the result of surface normals from least squares plane fitting by eliminating the effects of the noises and outliers. However, it also increased the run-time of the program with recalculating the surface normals.

First method of segmenting surface normals worked well for the dense and accurate point clouds and furthermore, was utilized to extract streets due to not varying the surface normals in the flat areas. The second method was carried out for the extraction of building roofs and worked well in most cases and especially applicable in case of low accurate point clouds. Second method was preferred to the first method in the extraction of building roofs due to updating the plane parameters and derived better result.

Availability of the 3D coordinates for each image pixel in the form of X-Y-Z image in addition to intensity values from RGB image or grayscale image can assist us to better interpret and recognize the objects and it can be considered the good combination in order to reach the goal of automatic 3D city reconstruction and object recognition.

Vegetation extraction has been led to failure in some image pixels that their intensity values differed too much from their realistic color in the nature. In addition, combination of X-Y-Z image with RGB image results in extraction of building roofs with more detail due to comparing three channels of the RGB image together and therefore makes it less flexible in region growing procedure in image space. Hence, the extracted building roofs boundaries from RGB image was not as good as building roofs boundaries from grayscale image.

Delineation of building roofs boundaries using 2D alpha shapes, result in better and realistic building roofs boundaries with more detail. However, it is performed by installing the huge library that increases the run-time of the program. Furthermore, there is trade-off between alpha values to choose the appropriate value.

The novelty of this work comparing to previous work is proceeding of surface growing based-approach with unlimited neighboring points and thereof no need of merging surface patches. In addition, robustness of the computation of surface normals assists us to discard outliers. Furthermore, segmentation in object space is applied in two different methods to increase the run time of the program in the procedure of extracting building roofs and streets. Furthermore, segmentation was investigated in three case studies to observe the effect of low accurate point clouds in the procedure of object recognition. In addition, RGB image compared to grayscale image in the process of segmentation in the image space to obtain better results.

As a recommendation for the future work, performing the segmentation procedure in object space and image space simultaneously by the usage of 3D point clouds in the form of X-Y-Z image in company of gray value from grayscale image may lead to better results of the segmentation and reach closer to the goal of automatic 3D city modeling and object recognition. In addition, performing segmentation in object space and thereafter in image space was also successful. However, it will derive better results by computing intersection of the adjacent building roofs for those building roofs that merged due to having similar intensity values. It is carried out by the usage of surface normals of the best fitted plane for each building roof and finding intersection lines between them.

## ACKNOWLEDGEMENTS

Ifp Institute of University of Stuttgart is gratefully acknowledged for providing the test data sets. Especially Prof. Dr.-Ing. Norbert Haala, Dipl.-Ing. Michael Peter and Dipl.-Ing. Mathias Rothermel are thanked for their advices and supports.

## REFERENCES

- [1] Jarzabek M., (2009) Segmentation and presentation of multiple 3D point clouds. Master thesis, University of Stuttgart, Germany.
- [2] Rabbani T., Heuvel F.A. van den, Vosselman G., (2006) Segmentation of point clouds using smoothness constraint. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 36(Part 5): 248-253
- [3] Vosselman G., Gorte B., Sithole G., Rabbani T., (2004) Recognizing structure in laser scanner point clouds. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 46(part 8/W2): 33-38
- [4] Bab-Hadiashar A. and Gheissari N., (2006) Range image segmentation using surface selection criterion, *IEEE transactions on image processing*, Vol. 15, No. 7, July 2006
- [5] Tovari D., (2006) Segmentation based classification of airborne laser scanner data. PhD Thesis, University of Karlsruhe, Germany
- [6] Rabbani T., (2006) Automatic reconstruction of industrial installations using point clouds and images. *Publications on Geodesy*, 62, Delft, The Netherlands
- [7] Filin S., (2002) Surface clustering from airborne laser scanning data. In *International Archives of Photogrammetry and Remote Sensing*, Vol. XXXII, 3A, pages 119–124, Graz, Austria.
- [8] Hoffman D. A. and Jain A. K., (1987) Segmentation and classification of range images, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 608-620, September 1987.
- [9] Filin S., Pfeifer N., (2005) Segmentation of airborne laser scanning data using a slope adaptive neighborhood. *ISPRS Journal of Photogrammetry and Remote Sensing*, 60, pp. 71-80
- [10] Marshall D., Lukacs G. & Martin R., (2001) Robust segmentation of primitives from range data in the presence of geometric degeneracy, *IEEE Transaction on Pattern Analysis and Machine Intelligence* 23(3): 304–314.
- [11] Klasing K., (2009) Surface-based Segmentation of 3D Range Data, Technical Report TR-LSR-2009-10-1, Institute of Automatic Control Engineering, Technische Universität München, Germany.
- [12] Biosca J.M., Lerma J.L., (2008) Unsupervised robust planar segmentation of terrestrial laser scanner point cloud based on fuzzy clustering methods. *ISPRS Journal of Photogrammetry and Remote Sensing*, 63, pp. 84-98
- [13] Jiang X.Y., Bunke H., (1994) Fast Segmentation of Range Images into Planar Regions by Scan Line Grouping. *Machine Vision and Applications* 7 (2): 115-122
- [14] Sithole G., Vosselman G., (2005) Filtering of airborne laser point scanner data based on segmented point clouds. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 36 (Part 3/W19): 66–71.
- [15] Han S.H., Lee J.H., Yu K.Y., (2007) An Approach for Segmentation of Airborne Laser Point Clouds Utilizing Scan-Line Characteristics. *ETRI Journal*, 29, pp. 641-648
- [16] Arman F. and Aggarwal J., (1993) Model-based object recognition in dense range images, *ACM Computing Surveys* 25, pp. 5–43, March 1993.
- [17] Hoover A., Jean-Baptiste G., Jiang X., Flynn P.J., Bunk, H., Goldgof D.B., Bowyer K., Eggert D.W., Fitzgibbon A., Fisher R.B., (1996) An Experimental Comparison of Range Image Segmentation Algorithms. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 18(No. 7): 673-689
- [18] Gorte B., (2002) Segmentation of TIN-structured surface models, *Symposium on Geospatial Theory, Processing and Applications; Working Group IV/6; Ottawa, Canada; July 8 – 12*
- [19] Fischer K., (2004) Introduction to Alpha Shapes.
- [20] Edelsbrunner H., Kirkpatrick D. G. and Seidel R., (1983) On the shape of a set of points in the plane. *IEEE Trans. Inform. Theory* IT-29, 551-559
- [21] Lee, I., Schenk, T., (2002) Perceptual organization of 3D surface points. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 34 (Part 3A): 193–198.

## Authors

**Mohammad Omidalizarandi** received a B.Sc. degree (Surveying Engineering) from the Iran University of Science and Technology (2004) and M.Sc. degree (Geomatics Engineering) from the University of Stuttgart (2011). He taught Photogrammetry, Image Processing and Geodesy for undergraduate students in Golestan University. His current research interests include Building reconstruction, Panorama, Image Matching and Videogrammetry.



**Mohammad Saadatseresht** received a B.Sc. degree (Surveying Engineering) (1996), M.Sc. degree (Photogrammetry) (1998) and PhD (Close range Photogrammetry) (2004) from Tehran University. He teaches Industrial Photogrammetry, Close range Photogrammetry, Photogrammetry and Image Processing for undergraduate, M.Sc. and PhD level Students. He is currently at the University of Tehran and working as associate professor.

