AUTOMATIC 3D RECONSTRUCTION OF BUILDINGS ROOF TOPS IN DENSELY URBANIZED AREAS

M. Gkeli [*], C. Ioannidis

Laboratory of Photogrammetry, School of Rural and Surveying Engineering, National Technical University of Athens, Greece
gkeli.maria1@gmail.com, cioannid@survey.ntua.gr

KEY WORDS: 3D reconstruction, RANSAC, plane fitting, moments of inertia, Automation, Building

ABSTRACT:

3D reconstruction of the urban environment constitutes a well-studied problem in the field of photogrammetry and computer vision, attracting the growing interest of the scientific community, for many years. Although the current state of the art present very impressive results, there is still room for improvements. The production of reliable and accurate 3D reconstructions is useful for a wide range of applications, such as urban planning, GIS, tax assessment, cadastre, insurance, 3D city modelling, etc. In this paper, a methodology for the automatic 3D reconstruction of buildings roof tops in densely urbanized areas, utilizing dense point clouds data, is proposed. It consists of three (3) main phases, each of which comprises a set of processing steps. In the first phase, the point cloud is simplified and smoothed. Outliers and non-roof elements are detected and removed utilizing shape, position and area criteria. In the second phase, the geometry buildings roof tops is optimized, by detecting and normalizing the edges. In the last phase, the reconstruction of the buildings roof tops is conducted. A progressive process, utilizing a plane fitting algorithm in combination with Screened Poisson Surface Reconstruction is performed. Buildings roof tops surfaces are produced and optimized. A software tool is developed and utilized for the implementation of the proposed methodology. The produced results are assessed and a comparison with another open-source software is conducted. The proposed methodology seems to be effective providing satisfactory results, as it can manage properly the really noisy point clouds of densely urbanized environments.

1. INTRODUCTION

Over the last twenty years, the automatic extraction and reconstruction of 3D buildings has been a major research focus, trying to replace the manual reconstruction of buildings from aerial imagery via stereoscopy or from LiDAR (Light Detection And Ranging) data, which are time consuming and laborious tasks. In order to achieve full automation, aerial imagery and LiDAR data need to be used in synergy in order to utilize their superior positional and height resolutions respectively. Recent advances in computer vision technology, the increasing quality of digital airborne cameras as well as the recent innovations in matching algorithms, have already demonstrated digital image matching as a valid alternative to airborne LiDAR. Therefore, the computation of dense 3D point clouds and the generation of Digital Surface Model (DSM) with surface resolution similar to the ground sampling distance of the available imagery, is feasible. 3D point clouds and DSMs, are of fundamental importance for 3D reconstruction of real-world, as they guide the overall reconstruction process, providing information about the studied surface. In recent literature several methods have been developed and proposed for building detection, recognition and reconstruction. Building outline consist an important geospatial information for several applications, such as urban planning, GIS (Geographic Information System), tax assessment, cadastre, insurance, 3D city modelling, etc (Wang, 2016). Automatic building extraction and reconstruction from remote sensing images are difficult tasks, as the detection accuracy and the quality of the produced surface, depends heavily on image resolution, quality and buildings shapes (Köhn et al., 2016).

In this paper, a methodology for the automated reconstruction of noisy buildings roof tops, is proposed. The main interest focuses on densely urbanized areas where buildings roof tops are distinguished by a high level of complexity due to the existence of non-roof, structural or non-structural man-mounted, elements. Based on some generic knowledge on building, it is possible to detect and remove such noise providing an effective reconstruction. As initial data for the function of the proposed procedure, a dense point cloud derived from nadir aerial images through dense matching techniques, is considered. In Section 2, a review of the current methods, techniques and algorithms, concerning the 3D reconstruction, is presented. In Section 3, the proposed methodology is analyzed. In Section 4, the developed software is presented and an implementation of the proposed procedure, is conducted. Furthermore, in Section 4, an assessment of the produced data and a comparison with the results of other software is made. Finally, in Section 5, the main conclusion referring to this work are presented.

2. RELATED WORK

3D reconstruction of real-world environment is an active research area. Although the current state of the art present very impressive results, there is still room for improvements. In recent literature several 3D reconstruction approaches have been proposed. These methods can be divided into three general categories based on the degree of contextual knowledge as: (a) Model-driven methods (parametric modeling), (b) Data-driven methods (non-parametric modeling), and (c) Hybrid methods (Gkeli et al., 2017). Model-driven or top-down approaches require primary knowledge about the shape of the buildings. The evaluation of the best-fitted 3D model is based on a library or grammar of predefined parametric shapes. Parametric methods are resilient to noisy and incomplete data, producing a topologically correct output model. However, their adaptability to various applications is limited by the narrow variety of the predefined shapes. In recent literature there are several Grammar-based automatic and semi-automatic reconstruction

procedures. The most well-known examples are Lsystems (Lindenmayer-systems), Shape Grammars (McKay et al., 2012; Karantzalos and Paragios, 2010), Split Grammar (Wonka et al. 2003), Computer Generated Architecture (CGA) Grammar (Müller et al., 2006), Formal Grammars (Becker and Haala, 2009) and Attributed Building Grammar (Yu et al., 2014; Yu et al., 2016).

Data-driven or bottom-up approaches are more flexible as they do not require any prior knowledge about a specific building structure. In recent literature, the majority of the proposed data-driven methods tend to extract points related to building's roof structure and classify them into different roof planes with 2D topology. The geometry of the roof can be described by a set of geometric primitives (planes, lines etc.) on which the 3D reconstruction procedure is based. Points can be clustered into planes based on similar attributes, such as: normal vectors, distance to a localized fitted plane or height similarities (Rottensteiner et al., 2014). Current data-driven methodologies and algorithms may be divided into four (4) general categories: (a) plane fitting based methods, (b) filtering and thresholding based methods, (c) segmentation based methods and (d) different supervised classification methods (Makantasis et al., 2015; Alidoost and Arefi, 2016). In the recent literature there are several approaches trying to apply plane fitting based methods on 3D point clouds, derived either from active sensors (e.g., LiDAR) or produced through photogrammetric procedures. The most well-known examples of algorithms used in this category are random sample consensus (RANSAC) algorithms (Fischler, 1981), least squares planar fitting algorithms (Omidalizarandi and Saadatseresgt, 2013) and plane fitting based algorithms (McClunea et al., 2016). Wang (2016), proposed a methodology for the detection and extraction of buildings roof outlines utilizing a dense point cloud derived from high-resolution aerial imagery. The proposed methodology tends to extract the ground surface using a polynomial surface adaptation method and then extract the buildings volumes by the production of nDSM (normalized DSM). Utilizing various radiometric and other criteria for the classification of all the elements located on buildings roof top, the outline of the roof is extracted through a split-and-merge method. Dal Poz and Fernandes (2016) proposed a methodology for the extraction of buildings boundaries and roof ridgelines, with the combined use of high-resolution aerial images and ALS (Aerial Laser Scanner) data. ALS data are utilized to limit the amount of straight lines representing the roof boundaries and then, Steger line detector and Canny edge detector are applied to the images, to identify lines within the limited area of the interior of the polyhedrons. Köhn et al. (2016), proposed a different method for the detection and reconstruction of the building roof, using aerial images. A line segment detector (LSD) is applied, in order to identify straight line segments through a region growing method among pixels with similar intensity and orientation. Utilizing some assumptions about buildings shape (rectangularity), buildings are detected. Finally, 3D buildings roofs are reconstructed by a RANSAC-based plane fitting procedure. An alternative and under-explored approach, is scan line segmentation which uses cross sections for segmenting planar features (Rottensteiner et al., 2014). McClune et al. (2014) proposed a methodology to derive the geometry of building boundaries using aerial images. The height differences along the 2D sections are examined using the DSM. Roof boundaries are identified as the parts with intense height differences.

Once the point cloud representing building's structure parts is defined through the abovementioned methods, the respective surface has to be reconstructed. Through numerous approaches, surface reconstruction methods can be divided into two (2) broad categories: (a) combinatorial algorithms and (b) implicit functions. Combinatorial algorithms introduce relationships between the points of the input data. These algorithms tend to divide the 3D space utilizing a tetrahedralization or voxel grid, clustering data through a topological analysis. The main disadvantage of these methods is that maintain the noise or corruption included in the initial data. An example of such an algorithm is the Ball pivoting algorithm (Bernardini et al., 1999). The basic idea of this algorithm is that a ball is moving around the 3D space trying to connect three points at a time. The ball is initially connected with two points of the input data and moving around until it intersects with a third point. Thus, triangles between the points of the input cloud are created. On the other hand, implicit functions present a more robust approach. In cases of noisy data, a common approach is to fit points using the zero set of an implicit function. Implicit methods usually based on the interpolation of the data or on approximating a surface near the input data. The most commonly known implicit methods of the latter case, are marching cubes (Lorensen and Cline, 1987) and Poisson surface reconstruction (Kazhdan, 2006; Kazhdan and Hoppe, 2013).
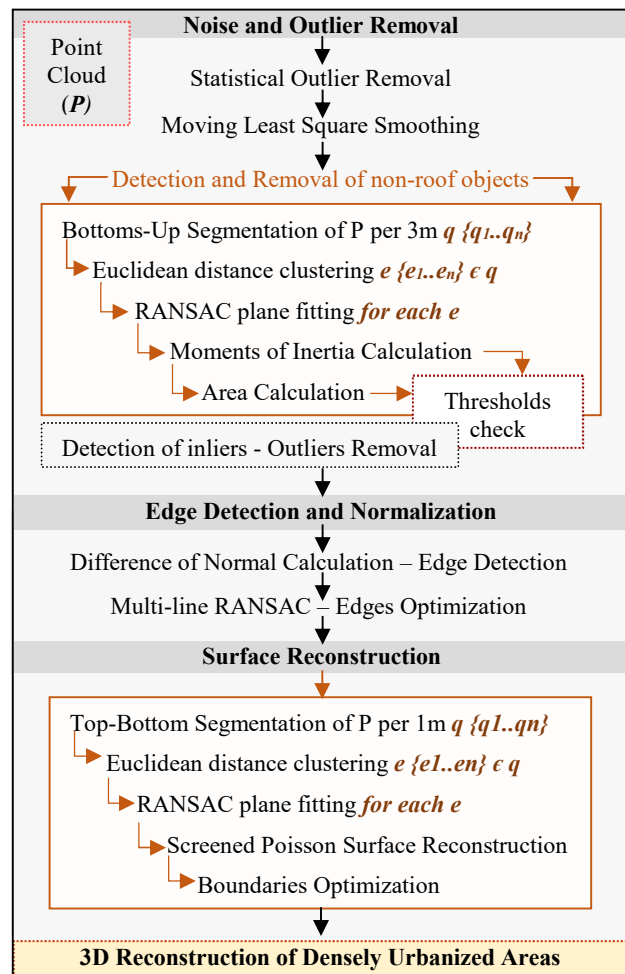

Figure 1. Workflow of the proposed methodology

Poisson surface reconstruction (Kazhdan, 2006) is a well-known technique, able to produce watertight surfaces from oriented point samples. While the most implementations

encountered in the current literature, use Poisson reconstruction in order to reconstruct surfaces from data derived by active sensors (e.g., terrestrial scanners) (Kazhdan, 2006; Kazhdan and Hoppe, 2013), this technique may provide very promising results utilizing other data sources (e.g., dense cloud from aerial images). Poisson surface reconstruction is resilient to noisy data and misregistration artifacts, and so it may be used for the reconstruction of complex urban scenes, following specific methodological steps. As it represents an implicit function, the produced result of this technique is a smooth approximate surface, positioned near the initial data. This encourages the usage of adaptive octree for the representation of the implicit function as well as for the solution of the Poisson system (Kazhdan, 2006). An important characteristic of Poisson surface reconstruction is the tendency to oversmooth the data. This may be undesirable in the most of the cases, but it can be adjusted to the current needs of each implementation, by defining the depth of the octree. The form of simple Poisson surface reconstruction is based on the normal field (inward pointing), which represents the boundary of a solid and may be interpreted as the gradient of the solid's indicator function. Thus, the main phases of Poisson surface reconstruction of a set of oriented points, may be defined as follows: (a) transformation of the oriented point cloud into a continuous vector field in 3D, (b) finding a scalar function whose gradients best match the vector field, and (c) extraction of the appropriate isosurface. In the proposed methodology screened Poisson reconstruction is utilized (Kazhdan and Hoppe, 2013). This approach is based on the traditional Poisson surface reconstruction (Kazhdan, 2006), incorporating positional constraints referred to the points data. By the term screening, a soft constraint that encourages the produced isosurface to pass through the input points, is implied. The main feature which differentiates the screened Poisson reconstruction over its traditional approach, is that the gradients are not constrained over the full 3D space, but the positional constraints are introduced only over the input points, which lie near a 2D manifold (Kazhdan and Hoppe, 2013).
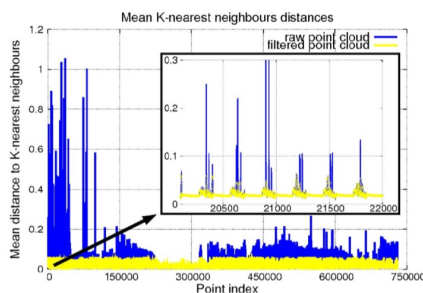

Figure 2. The mean distances to k = 30 neighbours before and after removal using α = 1 (Rusu, 2009)

3. METHODOLOGY

In this section, the proposed methodology for the automatic 3D reconstruction of buildings roof tops dense point cloud, in densely urbanized areas, is presented. It consists of three (3) main stages, as illustrated in Figure 1.

3.1 Point Cloud Noise and Outlier Removal

Before attempting to begin the reconstruction procedure and estimate the characteristics of a point in the initial point cloud, it is important to examine if this point represents adequately the underlying surface.

3.1.1 Statistical Outlier Removal (SOR): Outliers' removal may be conducted through the analysis of the surrounding neighbours of a point, utilizing mathematical analysis of their measured positions (Rusu, 2009). A point p_q may belong to the surface of an object if there are enough neighbours ($k \geq k_{min}$) in the vicinity of the query point. Certainly, point clouds densities vary accordingly to the current generation method and the particularities of the object's surface, such as shiny surfaces, depth discontinuities, or occlusions, creating false outliers and ghost structures. However, removing outliers is an essential step, leading to more reliable results and reduces the processing time.

The proposed methodology is based on a statistical analysis proposed by Rusu (2009). The algorithm examines each point of the point cloud ($p_q \in P$), calculating the mean distance d of its k neighbours. Then, a distribution over the mean distance space for the entire point cloud P is assembled and its mean μ_k and standard deviation σ_k are estimated. The main purpose is the production of a homogeneous point cloud, by excluding the points whose mean distance d differs greatest from the rest of the point of P (Figure 2). Thus, the remaining point cloud P^* derived according to equation (1). In the proposed methodology the selected k value is equal to 50 and α is equal to 1.0.

$$P^* = \{p^*_q \in P | (\mu_k - \alpha \cdot \sigma_k) \leq \overline{d}^* \leq (\mu_k + \alpha \cdot \sigma_k) \} \qquad (1)$$

where   P = initial point cloud
P^* = filtered point cloud
μ_k = mean deviation
σ_k = standard deviation
α = desired density restrictiveness factor
$\overline{d}$ = mean distance

3.1.2 Moving Least Square (MLS) smoothing: In this section, another optimization step able to filter out the noise and produce a smoothed surface, is presented. Moving least square algorithm (MLS) (Rusu et al., 2008; Alexa et al., 2003) tend to suppress outliers of a point cloud P by resampling (either upsampling or downsampling) and discarding unwanted data. Given a set of points P, MLS algorithm reproduce the complete smooth surface by fitting higher order bivariate polynomials to point in the vicinity of each query point p_q. In contrast to other interpolation or resampling techniques, MLS algorithm produce more robust results as the fitted surface passes through the original data. In the first step, the coordinates of each point p_q are normalized, utilizing the diagonal bounding box of the point cloud, ensuring a uniform distance between the points. The weight factor is estimated through equation (2):

$$h = \mu_d + k * \sigma_d \qquad (2)$$

where h = weight factor
μ_d = mean deviation of the distribution of mean distances between points
σ_d = standard deviation of the distribution of mean distances between points
k = nearest neighbours of the query point p_q

Next, the initial point cloud is resampled by estimating an approximate set of points Q, which is presented as a set of equidistant grid points, located in vicinity of the initial point cloud P. These points will be projected onto a local reference plane fitted through their k nearest neighbours P^k. The points in the local neighbourhood of each query point p_q may be

identified, either by the number k of their nearest neighbours or by using a fixed radius r, including a sub-set of P. Then, each point of Q is fitted to a surface which approximates the original data using a bivariate polynomial height function in a local Darboux frame. In most cases, 2nd order polynomials produce good results as the most areas are either planar or curved in only one direction. In the proposed methodology where the buildings roof tops may be estimated approximately through planar surfaces, the utilization of such a MLS algorithm with 2nd order polynomial function is selected. In this approach a MLS algorithm proposed by Rusu (2009) is selected. This algorithm is utilized in order to simply resample data without upsampling or downsampling the initial point cloud. Thus, the MLS algorithm optimize data position with respect to the approximate underlying smooth surface. The produced smoothed point cloud contains the same number of points with the initial cloud, possibly moved to a better approximate position if is needed.

3.1.3   Detection and Removal of non-roof objects: In cases of buildings roof tops there is a more complicated kind of noise referred to non-building's structure man-mounted objects (e.g., chimneys) or building's parts which are not included on the building roof top but to another lower height level from it (e.g., balconies). All these parameters of noise, strongly affect the overall reconstruction pipeline, resulting to inconsistent representations of the roofs structure. The detection and removal of such elements is essential. For the automatic extraction and reconstruction of buildings, some generic knowledge on buildings should be considered and utilized as geometric constraints. Building or floor height, size, shape regularity, roof surface smoothness, homogeneity and occlusions by trees or shadows, are some of the geometric and radiometric contextual properties about buildings structure and errors sources (Wang, 2016). In cases of densely urbanized areas, where buildings are partially-attached to neighbouring buildings with varying heights, the detection and removal of non-roof objects requires a bottom-up analysis of the building structure with respect to adequate geometric constraints.

The proposed methodology follows a sequence of five (5) discrete steps, trying to remove non-roof elements. The first step includes a bottoms-up segmentation of the point cloud. The algorithm searches the point cloud in order to find the point with the lower height and continues the segmentation by grouping points whose height difference is less than 3m. Given the fact that the initial point cloud may include points representing building's parts of a lower level, the 3m threshold is chosen as it corresponds to a typical floor height. This assumption divides the search space into smaller parts, facilitating the imposition of geometric constraints. Then, each of the height clusters derived from the first step, is processed separately. Each one of the height clusters include points with height difference less than 3 m. However, their distribution through the horizontal plane defined by X and Y axis, varies. Thus, for each one of the height clusters a simple Euclidean distance threshold based clustering algorithm (Rusu, 2009) is applied with a distance tolerance of 1m, in order to identify the individual points groups. At this phase, the point cloud is segmented in individual parts, each of which represents an element or object of buildings surfaces. Then, the process continues on the basis of certain assumptions. Given the fact that the surface of each building consists of several planar segments, a RANSAC based plane fitting algorithm (Fischler, 1981) with a threshold of 0.5m is applied to each one of the grouped points of the produced height clusters. The threshold of

0.5m is selected, in order to optimize the plane detection procedure, removing non-roof objects (noise) with height less than 0.5m. Subsequently, an initial normalization of the points position according to the detected planar segments is conducted, through the projection of the inliers on the detected plane and the removal of the outliers from the point cloud.

In the next step, each one of the detected groups which include the remaining points after the normalization of the point cloud, is checked against two certain criteria. The first criterion refers to a shape constrain. A group of points corresponding to a lower height level from the roof, in most of the cases, represent a balcony. This building element consist a non-roof structure, causing errors to the reconstruction procedure. However, it has some structural and geometric characteristics able to differentiate it from the rest building's elements. The basic dimensions of a balcony, length and width, are disproportionate resulting to a high valued ratio. Thus, this finding is used as a criterion for the detection of such elongated features in the point cloud and remove them. Subsequently, for each one of the groups of points, we calculate and construct the respective oriented bounding box (OBB), based on eccentricity and moment of inertia (Pratt, 2001; MOI, 2018). First of all, the covariance matrix is calculated and its eigen values and vectors are extracted. An iteration process begins where the major eigen vector is rotated continuously, around the other eigen vectors. For every current axis moment of inertia and eccentricity is calculated. Then, eccentricity is calculated for the obtained projection. In order to find the dimensions of the OBB, a transformation into the local frame of the box is calculated, keeping track of the minimum and maximum coordinates in each direction. Thus, utilizing the coordinates of OBBs corners, the calculation of its minimum and maximum sides is conducted. As threshold for the acceptance of a group of points as roof part, the ratio of 4 is selected, as it provides the best detection performance. The definition of the ratio is calculated through the equation (3).

$$\text{ratio}_{OBB} = \frac{Dmax_{OBB}}{Dmin_{OBB}} \leq 4 \tag{3}$$

where ratio_{OBB} = ratio between the max and min OBB distances
        $Dmax_{OBB}$ = maximum length of OBB
        $Dmin_{OBB}$ = minimum length of OBB

The second criterion refers to an area constrain. An object may be included in the main structure of a building only if its area exceeds a certain threshold. Otherwise, this object consists a non-structural object (e.g., chimneys) causing noise, and thus need to be excluded from the point cloud. As minimum accepted area of a group of point, $10m^2$ is selected. In reality, the size of $10m^2$ refer to enough small objects which may be included in the main structure of the building. However, the calculated area concerns a surface with multiple anomalies, and thus the minimum area assumption consists a satisfactory threshold for the proposed methodology.

3.2 Edge Detection and Normalization

One of the basic structural features of buildings roof tops is the regularity of their shapes. The roof tops edges are smooth and in the majority of cases their intersections are orthogonal. The normalization of edge points is of great importance, as the edges in the point cloud presented as "trembling lines" due to the non-uniform distribution of points, along with the remaining noise in the data (e.g., due to occlusions). In the proposed

methodology, edge detection is conducted with the multi-scale difference of normals (DoN) operator (Ioannou et al., 2012). DoN operator, is similar to the Difference of Gaussians (DoG) operator, providing very good results in identifying points saliency according to scale. In order to identify edges points, DoN operator utilize the estimated surface normal map of the point cloud. As the surface normals estimated at any radius r, reflect the underlying geometry of the surface at the scale of each support radius, this method is appropriate in order to identify areas with high differences which present the edges areas. In the literature there are several approaches arguing that edge detection may be conducted based on the surface curvature changes, but this is not the only way. While DoN seems to be enough simple, is shown to be surprisingly powerful as normals are less effected by noise compared to higher order derivative quantities such as curvature (Ioannou et al., 2012). In the proposed methodology normals estimation is conducted utilizing the approach proposed by Alexa et al. (2003) and extended by Rusu et al.(2008). According to this specific approach, normals estimation is conducted according to the characteristics of the points which are included in a support radius r around the query point p. This radius determines the scale in the surface structure which the normal represents. The basic idea of the DoN operator, is to compare the responses of the surface normals between two different radius $r_1 < r_2$. DoN operator Δn is calculated for each point p of the point cloud P. The utilization of the surface normals for a small and a large radius is selected in order to examine the upmost of the surface features. A small support radius is affected by small-scale surface structures. In contrast a large support radius is more robust in small-scale effects and represent the geometry of larger scale surface structures (Figure 3).
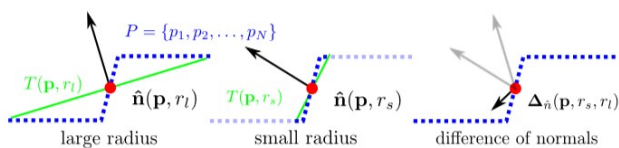

Figure 3. The meaning of large (left), small support radius (middle) and the calculated DoN operator (right) (Ioannou et al., 2012)

DoN is calculated for each point p in the point cloud P, and then a vector map is created. As each DoN presents the normalized sum of two normal vectors, its magnitude is varying through the values [0,1]. Setting a certain threshold on this magnitude, the extraction of the points related to the edges are extracted. In the proposed methodology a small support radius of 1m and a large support radius of 10m are used. As DoN threshold a value of 0.80 is selected, as presented the best performance. After extracting the detected edges points, a RANSAC based line fitting algorithm (Fischler, 1981; Rusu, 2009) with a threshold of 0.5m, is imposed upon them. Thus, edges points are normalized, by projecting them onto the estimated line, in order to satisfy as much as possible, the regularity criterion of roofs structure.

3.3 Partial Surface Reconstruction

Reconstructing 3D surfaces is a well-studied problem, attracting the growing interest of the scientific community, for many years. Although there is an impressive amount of different approaches (Yu et al., 2014; Makantasis et al., 2015; Yu et al., 2016; Alidoost and Arefi, 2016; Köhn et al., 2016; McClunea et al., 2016), aiming the 3D reconstruction of the real world, there is still room for improvements. In the proposed methodology,

screened Poisson surface reconstruction (Kazhdan and Hoppe, 2013) is utilized, for the reconstruction of buildings roof tops in densely urbanized areas.

In this phase the 3D reconstruction of the refined point cloud derived from the previous steps, is attempted. The main idea of this algorithm is to provide a top-down approach, reconstructing partially the detected planar surfaces, constructing the buildings roof tops. Due to the high level noise presented in the data of an urban scene, as described in section 3.1.3, the reconstruction of the planar roof tops surfaces is not a trivial neither an easy target. The algorithmic procedure followed in this phase, consist of five (5) certain steps. In the first step, a top-down segmentation per 1m, of the refined points, is conducted. The algorithm searches the point cloud in order to find the highest point and continues the segmentation by grouping points whose height difference is less than 1m. The selection of 1m, as a threshold for the segmentation process, is decided based on the assumption that in each section of 1m of the vertical dimension of a building it is possible the existence either of a planar surface or a vertical surface. Due to the fact, that the proposed procedure refers to point clouds produced from nadir aerial images, points presenting the vertical elements of a building may be very few and presented rarely in the point cloud, causing noise. Thus, at each section of 1m the existence of a planar surface, is very likely. However, in contrast with the possible distribution of planar surfaces along the vertical direction, their horizontal distribution is not obeying the same premise. In the horizontal direction, buildings roof tops may be connected semi-detached or detached with the neighbouring buildings. Thus, in the next step, a simple Euclidean distance threshold based clustering algorithm is applied to each one of the height clusters, with a distance tolerance of 3m in order to identify the individual points groups, which represent neighbouring planar surfaces. Subsequently, a RANSAC based plane fitting algorithm with a threshold of 0.5m, is applied to each one of the grouped points of the produced height clusters. Subsequently, the detected inliers are projected on the estimated plane and the outliers are removed from the point cloud. This step is important, in order to optimize each one of the grouped points, reducing the remaining noise from the previous stages. The final acceptance of each point group as possible representation of a building's planar element, is done if the corresponding area exceeds 5m^2.

Once a group of points is optimized and checked against the threshold of 5m^2, the surface reconstruction of the corresponding planar surface is conducted, through utilizing screened Poisson surface reconstruction (Kazhdan and Hoppe, 2013), with octree depth equal to 8. The reconstruction of each one of the point groups separately, is selected, in order to force the produced surface to fit to these certain points without trying to change the final shape of the produced surface, in order to fit to the whole point cloud. The next step, concerns a particularity that presented from screened Poisson surface reconstruction algorithm. Poisson surface reconstruction algorithm tends to surround the data, in order to achieve the best possible fitting of the calculated surface. Thus, unnecessary surface extensions are created at the boundaries of the studied object. These extensions need to be removed, in order to produce a reliable reconstruction. For this reason, the concave's hull polygon (Rusu, 2009) referred to each one of the point groups, is calculated. Subsequently, the final reconstruction presenting the roof tops of the buildings in densely urban areas, is produced cropping Poisson Surface along the boundaries of Concave's Hull polygon with alpha parameter equal to 5.

4. IMPLEMENTATION

For the implementation of the proposed methodology a software in programming language C++ has been developed, utilizing the libraries of PCL - Point Cloud Library 1.8.0 (PCL, 2018), VTK - Visualization Toolkit (VTK, 2018) and Eigen (Eigen, 2018). The developed software requires as input an ASCII or Binary PCD (Point Cloud Data) file, containing the coordinates X, Y, Z of a dense point cloud representing an urban scene, and produce as output a VTK file, containing the reconstructed surface of the buildings roof tops. For the evaluation of the proposed methodology a test was conducted. The produced results are assessed and compared to the results of other free software.

The test area is located in the centre of Athens, Greece and is characterized by densely constructed semi-detached buildings with varying heights. As it is depicted in the Figure 4 the separation between the individual roof tops is quite difficult. Subsequently, the roof tops of the buildings are not represented by smooth surfaces, but on the contrary they have complex geometry due to the existence of several non-structural objects positioned on their surfaces. As test data for the implementation of the proposed procedure a dense point cloud derived by a set of aerial images has been utilized (Figure 5, left). The point cloud was inserted in the developed software, so as the 3D reconstruction of buildings roof tops be produced.

Figure 4. An orthophoto depicting the tested area of densely urbanized area in the centre of Athens, Greece

At the first step of the proposed methodology, outliers are removed from the initial point cloud and then an MLS smoothing with radius of 1m, is applied. Subsequently, a bottom to top segmentation of the point cloud is conducted, clustering the cloud into several height levels. As depicted in Figure 5, the greater amount of noise has been removed. Each of the levels presented with a random selected colour. In the step, Euclidean distance-based clustering is conducted, in order to divide the point cloud into several groups of point, each one of them consist a candidate element of building's structure. Each one of these groups are assessed against the shape and area criterions proposed, in section 3.1.3. In Figure 6, the results of this step are presented. Each one of the detected group of points is presented with different random colour, while the calculated oriented bounding boxes (OBB) on which the shape criterion is applied, are identified with white colour. The results of the first step are satisfying, as the majority of noise and non-roof objects were detected and removed from the point cloud. However, in some cases non-roof structures, such as balconies,

with area size more than 10m^2 located in a lower level from the roof tops, remain in the cloud and presented as isolated segments. An example of such false-detected roof element, is depicted in a red circle in Figure 6, left. Also, the existence of repetitive elongated structures on the surface of the roof top, may be confused and be considered as noise, so these structures are considered as noise and removed from the point cloud (Figure 7). However, the presentence of such particular examples is not constituted a common case. Thus, the detection and optimization of the point cloud, removing outliers and non-roof elements, leads to satisfactory results.
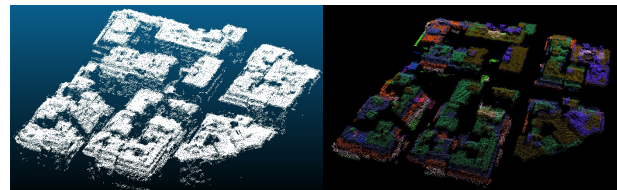

Figure 5. The input point cloud (left) and the segmented per 3m cleaned point cloud of the urban area (right)
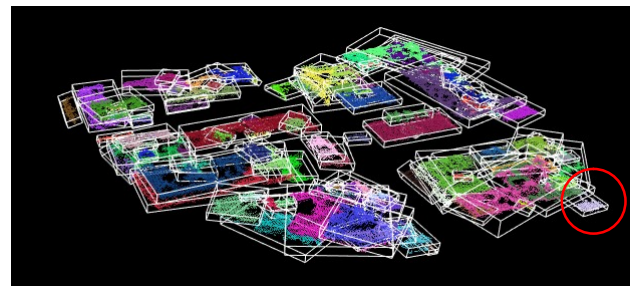

Figure 6. The detected roof elements presented in random colour, surrounded by the OBB and a false-detection presented in the red circle

In the next step DoN operator is calculated, and the detected edges of the point cloud are refined. In Figure 8, the detected edges (left) and the refined point cloud (right) are presented. The refined boundaries are normalized, in order to satisfy the shape regularity constrain. Finally, in the last step of the proposed procedure, the automatic surface reconstruction of the buildings roof tops is conducted. The produced results are presented in Figure 9. Although there were some false-detections referring to roof and non-roof elements, the produced results are satisfactory as the final model represents reliably the buildings roof tops, as smooth surfaces (Figure 9). The remaining noise is rarely presented as isolated surface segments far away from the buildings.

In the next step, the evaluation of the produced data against the corresponding reconstruction results derived from another open-source software, is conducted. For the comparison, the software of MeshLab (MeshLab, 2018) and CloudCompare (CloudCompare, 2018) were utilized. The test point cloud inserted in each one of the software. At the first phase the noise is removed from the data, and then point normals are calculated based on a neighbourhood of k = 30 neighbouring points. At the last phase, the Screened Surface reconstruction algorithm with a depth of 8, is performed. In Figure 10, the results derived from MeshLab are depicted. The building roof tops presented as noisy rough surfaces, while the roof tops boundaries are presents as wavy segments. Also, at sparse positions on the roof tops surfaces, holes are presented, changing their geometry. Non-roof elements remain in the scene complicated the processing procedures. Subsequently, the results derived from

CloudCompare are depicted in Figure 11. The produced results are presented smoother in contrast with the previous of MeshLab. However, roof tops surfaces consist of multiple anomalies while their boundaries are difficult to be identified accurately. Furthermore, noise referring to non-roof elements, as balconies, vertical walls, are remain in the data as well as holes on roof tops surfaces are appeared. Thus, according to the abovementioned findings and assessments, appears that the proposed methodology is effective, reliable, can cope with noisy and complex data providing a promising solution for the reconstruction of densely urban scenes.

Figure 7. An example of false-detection of noise. In the red circles, the removed roof top elements are presented in the orthophoto (left) and in the result of the first step of the proposed methodology (right)
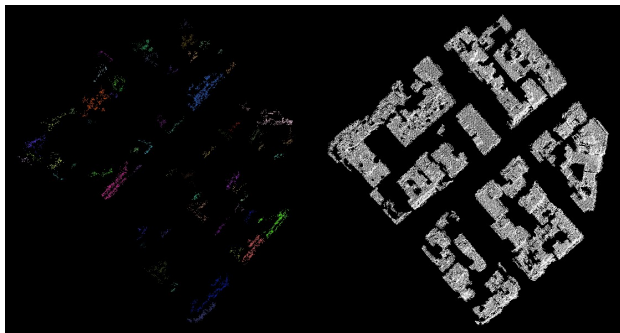

Figure 8. The detected edges (left) by DoN operator and the refined point cloud with normalized edges (right)
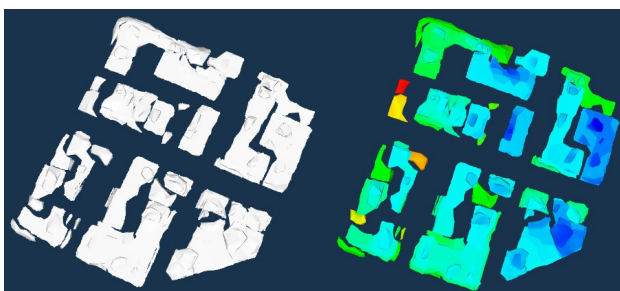

Figure 9. The result of the proposed 3D reconstruction methodology (left) and the height colorized representation (right)

5. CONCLUSIONS

This paper proposes a methodology for the automated reconstruction of point clouds representing complex and noisy urban areas. The proposed methodology is able to detect non-roof top building's elements and exclude them from the data. Simultaneously, by imposing several constrains referring to floor height, size, shape regularity, surface smoothness and homogeneity succeeds to produce a reliable and qualitative result, by presenting each building's roof top as a smooth and enough regular surface. The developed software, implementing

the proposed methodology, produce a remarkable result, which outperforms over the compared software. The data produced through the proposed methodology, may be utilized by a wide variety of applications, such as 3D Cadastre, urban planning, GIS etc., facilitating the necessary 3D reconstruction procedures.
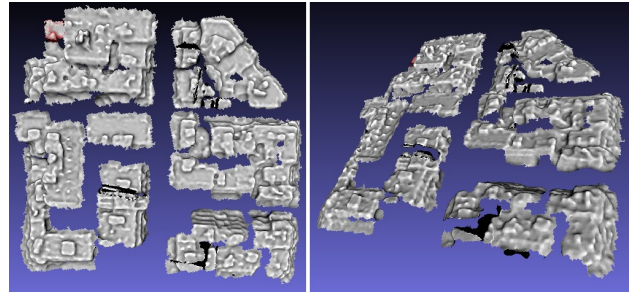

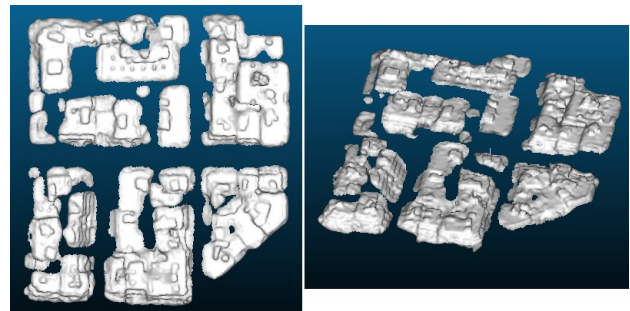Figure 10. Screened Poisson Reconstruction of the tested area in MeshLab

Figure 11. Screened Poisson Reconstruction of the tested area in CloudCompare

REFERENCES

Alexa, M., Behr, J., Cohen-Or, D., Fleishman, S., Levin, D., Silva, C.T., 2003. Computing and rendering point set surfaces. In: *IEEE Transactions on Visualization and Computer Graphics*, Vol. 9(1), pp. 3–15.

Alidoost, F., Arefi, H., 2016. Knowledge Based 3d Building Model Recognition Using Convolutional Neural Networks from LIDAR and Aerial Imageries. In: *ISPRS International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. XLI-B3, pp. 833-840.

Becker, S., Haala, N., 2009. Grammar supported facade reconstruction from mobile lidar mapping. In: *ISPRS Workshop, CMRT09-City Models, Roads and Traffic*, vol. 38, p. 13.

Bernardini, F., MIttleman, J., Rushmeier, H., Silva, C., Taubin, G., 1999. The ball-pivoting algorithm for surface reconstruction. In: *IEEE Transactions on Visualization and Computer Graphics,* Vol: 5(4), pp. 349–359.

Dal Poz, A.P., Fernandes, V.J.M., 2016. Extraction of Roof Lines from High-Resolution Images by a Grouping Method. In: *ISPRS International Archives of Photogrammetry, Remote*

Sensing and Spatial Information Sciences, Vol. XLI-B3, p.853-857.

Eigen, 2018. Eigen: C++ template library for linear algebra http://eigen.tuxfamily.org/index.php?title=Main_Page (March 2018).

Fischler, M.A., Bolles, R.C., 1981. Random sample consensus: a paradigm for model fitting with application to image analysis and automated cartography. *Commun. Assoc. Comp. Mach.*, Vol. 24, pp. 381–395.

Gkeli, M., Ioannidis, C., Potsiou, C., 2017. Review of the 3D Modelling Algorithms and Crowdsourcing Techniques - An Assessment of their Potential for 3D Cadastre. In: *FIG Working Week 2017 – ''Surveying the world of tomorrow – From digitalisation to augmented reality''*, Helsinki, Filand, pp. 1-23.

Ioannou, Y., Taati, B., Harrap, R., Greenspan, M., 2012. Difference of normals as a multi-scale operator in unorganized point clouds. In*: Proc. 2nd Int. Conf. 3D Imaging, Model., Process., Vis.Transmiss.*, pp. 501–508.

Karantzalos, K., Paragios, N., 2010. Large-scale building reconstruction through information fusion and 3-d priors. In: *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 48(5), pp. 2283-2296.

Kazhdan, M., Hoppe, H., 2013. Screened Poisson surface reconstruction. *ACM Trans Graph*, 32:29, doi: 10.1145/2487228.2487237.

Kazhdan, M., Bolitho, M., Hoppe, H., 2006. Poisson surface reconstruction. In: *Eurographics symposium on geometry processing*.

Köhn, A., Tian, J., Kurz, F., 2016. Automatic Building Extraction and Roof Reconstruction in 3k Imagery Based on Line Segments. In: *ISPRS International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. XLI-B3, pp. 625-631.

Lorensen, W., Cline, H., 1987. Marching cubes: A high resolution3dsurfacereconstructionalgorithm. *SIGGRAPH*, pp. 163–169.

Makantasis, K., Karantzalos, K., Doulamis, A., Loupos, K., 2015. Deep Learning-Based Man-Made Object Detection from Hyperspectral Data. In*: International Symposium on Visual Computing, Springer International Publishing*, pp. 717-727.
McClune, A.P., Miller P.E., Mills J.P., Holland D., 2014. Automatic urban 3D building reconstruction from multi-ray photogrammetry. In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol: XL-3, pp. 219-226.

McClunea, A.P., Millsa, J.P., Millerb, P.E., Hollandc, D.A., 2016. Automatic 3d Building Reconstruction from a Dense Image Matching Dataset. In: *ISPRS International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. XLI-B3, pp. 641-648.

McKay, A., Chase, S.C., Shea, K., Chau, H.H., 2012. Spatial grammar implementation: From theory to useable software. *AI EDAM (Artificial Intelligence for Engineering Design, Analysis and Manufacturing)*, Vol. 26, pp. 143–159.

MOI, 2018. Moment of inertia and eccentricity based descriptors. Point Cloud Library (PCL) http://pointclouds.org/blog/hrcs/sushakov/index.php (April 2018).

Müller, P., Wonka, P., Haegler, S., Ulmer, A., van Gool, L., 2006. Procedural modeling of buildings. In: *ACM Transactions on Graphics (Tog)*, vol. 25(3), pp. 614-623, doi:10.1145/1179352.1141931. 3.

Omidalizarandi, M., Saadatseresgt, M., 2013. Segmentation and Calssification of Point Clouds from Dense Aerial Image Matching. In: *The International Journal of Multimedia & Its Applications*, Vol. 5(4), pp. 33-50.

PCL, 2018. Point Cloud Library (PCL) http://pointclouds.org/ (April 2018).

Pratt, W., 2001. Digital Image Processing. 3rd edition, John Wiley & Sons, New York.

Rottensteiner, F., Sohn, G., Gerke, M., Wegner, J.D., Breitkopf, U., Jung, J., 2014. Results of the ISPRS benchmark on urban object detection and 3D building reconstruction. In: *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 93, pp. 256-271.

Rusu, R.B., 2009. Semantic 3D Object Maps for Everyday Manipulation in Human Living Environments. PhD thesis, Computer Science department, Technische Universitaet Muenchen, Germany.

Rusu, R.B., Marton, Z.C., Blodow, N., Beetz, M., 2008. Persistent point feature histograms for 3d point clouds. *Autonomous systems,* Vol. 10, p. 119.

VTK, 2018. The Visualization Toolkit (VTK) https://www.vtk.org/ (April 2018).

Wang, J., Wang, W., Li, X., Cao, Z., Zhu, H., Li, M., He, B., Zhao, Z., 2016. Line Matching Algorithm for Aerial Image Combining image and object space similarity constraints. In: *ISPRS International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. XLI-B3, pp. 783-788.

Wonka, P., Wimmer, M., Sillion, F., Ribarsky, W., 2003. Instant architecture. *ACM*, Vol. 22(3), pp. 669-677.

Yu, Q., Helmholz, P., Belton, D., 2016. Evaluation of Model Recognition for Grammar-Based Automatic 3d Building Model Reconstruction. In: *ISPRS International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. XLI-B4, pp. 63-69.

Yu, Q., Helmholz, P., Belton, D., West, G., 2014. Grammar-based Automatic 3D Model Reconstruction from Terrestrial Laser Scanning Data. In: *ISPRS International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. XL-4, pp. 335-340.