ISPRS Journal of Photogrammetry and Remote Sensing 94 (2014) 183–193

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

journal homepage: www.elsevier.com/locate/isprsjprs

A global optimization approach to roof segmentation from airborne lidar point clouds

Jixing Yan [a], Jie Shan [a,b,*], Wanshou Jiang [c]

[a] School of Remote Sensing and Information Engineering, Wuhan University, 129 Luoyu Rd., 430079 Wuhan, China
[b] School of Civil Engineering, Purdue University, West Lafayette, IN 47907, USA
[c] State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, 129 Luoyu Rd., 430079 Wuhan, China

ARTICLE INFO

Article history:
Received 6 August 2013
Received in revised form 27 April 2014
Accepted 29 April 2014
Available online 4 June 2014

Keywords:
Segmentation
City modeling Buildings
Lidar
Point clouds
Global optimization

ABSTRACT

This paper presents a global plane fitting approach for roof segmentation from lidar point clouds. Starting with a conventional plane fitting approach (e.g., plane fitting based on region growing), an initial segmentation is first derived from roof lidar points. Such initial segmentation is then optimized by minimizing a global energy function consisting of the distances of lidar points to initial planes (labels), spatial smoothness between data points, and the number of planes. As a global solution, the proposed approach can determine multiple roof planes simultaneously. Two lidar data sets of Indianapolis (USA) and Vaihingen (Germany) are used in the study. Experimental results show that the completeness and correctness are increased from 80.1% to 92.3%, and 93.0% to 100%, respectively; and the detection cross-lap rate and reference cross-lap rate are reduced from 11.9% to 2.2%, and 24.6% to 5.8%, respectively. As a result, the incorrect segmentation that often occurs at plane transitions is satisfactorily resolved; and the topological consistency among segmented planes is correctly retained even for complex roof structures.

1. Introduction

Building roof reconstruction is of a current research interest in 3D city modeling. Because of being able to directly collect dense, accurate 3D point clouds of urban objects, lidar (Light Detection and Ranging) technology provides an efficient solution to this need. Reported methods for building roof reconstruction mostly fall into two categories: data-driven (bottom–up) and model-driven (top–down). In terms of data-driven methods, a common assumption is that a building is a polyhedron consisting of planes and edges. As a crucial step, the point clouds of a building are usually segmented into disjointed planar regions. Subsequent tasks, including roof edge extraction and topologic reconstruction, are dependent on the quality of segmentation. A poor segmentation may make these tasks fail. As for model-driven methods, a building is assumed to be an assembly of roof primitives (e.g., gable roof and hipped roof), which and whose topology are predefined in a model library (Tarsha-Kurdi et al., 2007a; Huang et al., 2013). However, roof segmentation is still a required step in many model-driven methods, such as the graph matching approach (Verma et al., 2006; Oude Elberink and Vosselman, 2009). A poor segmentation may alter the topology among roof planes and make the matching task fail (Oude Elberink and Vosselman, 2009).

Building roof segmentation can be accomplished via various approaches, such as data clustering, region growing, energy minimization, and model fitting. A review of these approaches can be found in (Awwad et al., 2010; Sampath and Shan, 2010). Data clustering is basically a statistical technique that classifies the point clouds into primitives based on certain pre-calculated local surface properties or features. Filin and Pfeifer (2006) propose a slope adaptive neighborhood for such calculation. Considering the variations in local point density, Sampath and Shan (2010) use the Voronoi neighborhood to estimate the local surface properties, whereas Lari et al. (2011) use a cylindrical neighborhood for this purpose. As for clustering the feature vectors, mode-seeking (Filin and Pfeifer, 2006), conventional mean-shift (Melzer, 2007), and fuzzy k-means (Sampath and Shan, 2010) are applied. In spite of the popularity and efficiency of this approach, it suffers the difficulty in neighborhood definition and is sensitive to noise and outliers.

Region growing is a region-based segmentation method that partitions point clouds into disjoint homogenous regions. It starts with a selected seed point or region and expands to its neighboring points. Gorte (2002) selects the triangles in triangulated irregular

* Corresponding author. Address: Jie Shan School of Civil Engineering, Purdue University, West Lafayette, IN 47907, USA.
E-mail addresses: yanjx@whu.edu.cn (J. Yan), jshan@ecn.purdue.edu (J. Shan), jws@whu.edu.cn (W. Jiang).

http://dx.doi.org/10.1016/j.isprsjprs.2014.04.022
0924-2716/

networks (TINs) as seed regions and extends them to neighboring triangles. Zhang et al. (2006) perform a local plane fitting at points and select the points with good planarity as seed points. To obtain robust seed points, Chauve et al. (2010) develop an iterative Principal Component Analysis (PCA) to estimate local planarity. You and Lin (2011) present a non-iterative approach using tensor voting for this purpose. Unlike the aforementioned approaches, Dorninger and Pfeifer (2008) determine seed clusters (regions) by a hierarchical clustering approach. As for expanding seed regions, similarity measures such as distances of points to planes (Zhang et al., 2006; Dorninger and Pfeifer, 2008; Chauve et al., 2010) and angle differences between normal vectors (Dorninger and Pfeifer, 2008; Chauve et al., 2010; You and Lin, 2011) are used. Nevertheless, region growing is susceptible to the selection of seed regions (Awwad et al., 2010) and difficult to stop when transitions between two regions are smooth (Sampath and Shan, 2010).

The energy minimization approach is a global solution that formulates the segmentation as an optimization problem. Its objective function may consist of fidelity to data, continuity of feature values and compactness of segment boundaries (Kim and Shan, 2011; Vitti, 2012). A widespread application of this approach to image segmentation can be found in (Vitti, 2012). As for the segmentation of lidar data, multiphase level set approach is adopted to segment planar roof primitives under an energy minimization formulation (Kim and Shan, 2011). Compared to the RANSAC (RANdom SAmple Consensus, Fischler and Bolles, 1981) based approaches, it is global and multiple roof planes can be segmented at one time. However, a common shortcoming of this approach is that poor segmentation may occur when the energy function converges to a local minimum.

Since the reconstructed models are dependent on the robust estimate at planar primitives, robust model fitting methods such as RANSAC and Hough transform (Duda and Hart, 1972) are also applied to roof segmentation. Lidar points that fit a mathematical plane with most inliers are first extracted and regarded as a planar segment. This approach is robust to noise and outliers, but it tends to result in spurious planes (Tarsha-Kurdi et al., 2007b; Yan et al., 2012). With the help of available building ground plans, Vosselman and Dijkman (2001) split the dataset into small parts before applying Hough transform to prevent the detection of spurious planes. Some extended RANSAC considering local surface normals (Bretar and Roux, 2005; Schnabel et al., 2007; Awwad et al., 2010; Chen et al., 2012) are also developed for this purpose. Considering spatial connectivity, Zhang et al. (2006) and Chauve et al. (2010) combine model fitting and region growing. Nevertheless, most of the model fitting approaches are order-dependent and based on a single-model. Segmentation results are dependent on the order in which the planes are extracted. When multiple planes are present, each plane instance needs to be sequentially extracted. As a result, points at transitions between roof faces will be assigned to the first extracted planes. In most cases, this approach performs well with some additional constraints. However, for complex roof structures it tends to result in mistakes, such as spurious planes (segments that do not exist in reality), over-segmented planes (one actual plane is segmented into multiple segments), and under-segmented planes (multiple actual planes are segmented into one segment).

To overcome these problems, this paper seeks a global optimization solution to the problem of roof segmentation from airborne lidar point clouds. A multi-label (plane) optimization approach is introduced for this purpose. It intends to reduce the number of mistakes derived from a plane fitting based on region growing and to improve the topological consistency among segmented planes. In our study, the point clouds of building roofs are first extracted from airborne lidar data. Starting with initial planes derived from a plane fitting approach, a global energy function consisting of fidelity to data, spatial smoothness, and the number of models (i.e., the number of planes) is constructed to optimize the segmented planes. Comparing to existing approaches, the proposed method incorporates spatial smoothness between data points into plane fitting. It can produce spatially coherent segments and improve the segmentation quality. Additionally, the proposed approach is global, i.e., multiple roof planes are determined at the same time and their corresponding model parameters can be refined when minimizing the energy function. It yields both high completeness and high correctness rates. More noticeably, the incorrect segmentation that often occurs at plane transitions is satisfactorily avoided and the topological consistency among segmented planes is correctly retained.

The remainder of the paper is structured as follows. Section 2 formulates the segmentation task as a multi-label (plane) optimization problem and presents a graph cuts based global solution. Section 3 starts with the test data description, followed by a presentation of individual and overall test results to demonstrate the solution procedure. Assessment and discussion constitute Section 4, where we define our quality metrics, examine the sensitivity of the solution to relevant parameters, and assess the metric quality and topologic quality of the segmentation outcome. Both quantitative and qualitative evaluations are presented. Section 5 consists of our concluding remarks on the properties of the method and future effort.

2. Multi-label optimization

2.1. Formulation

The segmentation task can be noted as a labeling problem and formulated in terms of energy minimization. Eq. (1) provides such a formulation (Delong et al., 2012; Isack and Boykov, 2012)

$$E(L) = \overbrace{\sum_{p\in P} D_p(L_p)}^{data\ cost} + \overbrace{\sum_{p,q\in N} w_{pq}\cdot\delta(L_p\neq L_q)}^{smooth\ cost} + \overbrace{\sum_{l\in\mathcal{L}} h_l\cdot\delta(l\in\mathcal{L}')}^{label\ cost} \quad (1)$$

where $\mathcal{L}$ is a given set of labels (planes) and $\delta(.)$ is an indicator function. Let P represent a set of data points, the multiple labeling task is to assign a point $p \in P$ a label $L_p \in \mathcal{L}$ such that the labeling L minimizes the energy $E(L)$, where $\mathcal{L}'$ is the set of labels appearing in L and N is an assumed neighborhood for data points. Three energy terms are considered in the energy formula. The data cost term (the first term in Eq. (1)) measures the discrepancy between data points and labels. It is the sum of the distances of points to their assigned labels. The smooth cost term (the second term in Eq. (1)) measures the label inconsistency between neighboring points. It is the sum of weight w_{pq} of each pair of neighboring points p and q that are assigned to different labels. The label cost term (the third term in Eq. (1)) measures the number of labels appearing in L. It is the sum of the label cost h_l of each label $l \in \mathcal{L}'$. Fig. 1 illustrates a labeling of data points and their fitted lines. Two lines A and B are fitted to

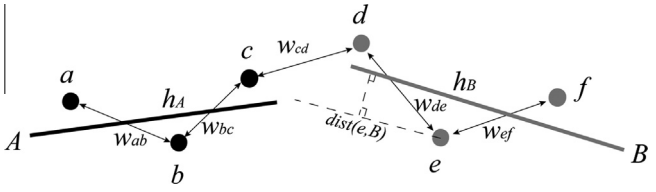


Fig. 1. Labeling of data points and their fitted lines. The double-arrowed lines link each pair of neighboring points. Data points and their corresponding fitted lines are shown with the same shade.

data points and the value of label costs is $H_A + H_B$. As a result, the data cost can be calculated as $dist(a,A)+dist(b,A)+dist(c,A)+dist(d,B)+dist(e,B)+dist(f,B)$ and the value of smooth costs is w_{cd}, where $dist(.)$ is the perpendicular distance of a point to its fitted line.

2.2. Cost terms

Since a plane corresponds to a unique label, the multi-plane optimization can be stated as a multi-label optimization problem. Accordingly, the three terms in Eq. (1) can be adapted for our task. This section will discuss how to construct the energy function for multi-plane optimization.

The data cost term penalizes the disagreement between a point and its assigned label. In the paper, the perpendicular distances of lidar points to the segmented planes is used to measure the disagreement. For a plane $ax + by + cz + d = 0$ $(a^2 + b^2 + c^2 = 1)$, its paramter vector is (a, b, c, d). Assuming that the plane fitting errors of inliers follow the Gaussian distribution, the data cost between a lidar point p (x_p, y_p, z_p) and its assigned label (plane) $L_p(a_p, b_p, c_p, d_p)$ is calculated as

$$Dp(Lp) = -\ln\left(\frac{1}{\sqrt{2\pi}\Delta d}\cdot\exp\left(-\frac{dist(p,Lp)^2}{2\Delta d^2}\right)\right) \quad (2)$$

where

$$dist(p,L_p) = \begin{cases} ax_p + by_p + cz_p + d_p & L_p \neq L_{outlier} \\ 2\Delta d & L_p = L_{outlier} \end{cases} \quad (3)$$

In the aforementioned equation, $L_{outlier}$ is an extra label for noise or outliers and Δd is a distance threshold in plane fitting. Notably, the distance between lidar points and label $L_{outlier}$ is a constant value, which is set to $2\Delta d$ in this paper. It means that the lidar points with distances to their corresponding planes larger than $2\Delta d$ are more likely to be noise or outliers.

The smooth cost term penalizes the label inconsistency between neighboring points. To minimize the energy function, neighboring points are encouraged to have similar labels (Delong et al., 2012). For this purpose, the neighborhood from derived Triangulated Irregular Networks (TINs) is used. If a pair of neighboring lidar points p and q fit the same plane, the smooth cost between them is 0; otherwise, the smooth cost is 1. The smooth cost term considers the discontinuity at plane transition under a global fitting and produces spatially consistent segments. Since adjacent points are more likely to fit the same model, w_{pq} is set inversely proportional to the distance between points (Isack and Boykov, 2012), i.e.

$$w_{pq} = \exp(-\|\mathbf{p} - \mathbf{q}\|) \quad (4)$$

The label cost term penalizes the number of labels. Fewer labels are encouraged to be used to represent data compactly (Delong et al., 2012). The term is introduced to reduce the number of redundant planes in the segmentation results. When minimizing the energy function, if the costs (data costs and smooth costs) of eliminating a label l are smaller than h_l, then the label l will be eliminated or merged with other labels. A large label cost can help to eliminate redundant planes, but small segmented planes may be missed after the label optimization. To ensure that small roof planes are kept after the label optimization, h_l is written as

$$h_l = \begin{cases} n \times \left(\frac{1}{2} - \ln\frac{1}{\sqrt{2\pi}\Delta d}\right) & L_p \neq L_{outlier} \\ 0 & L_p = L_{outlier} \end{cases} \quad (5)$$

where n is the minimum number of lidar points required for a valid plane. Except for the labels for noise or outliers, every label has the same label cost. However, the label costs derived from Eq. (5) may be smaller for some large redundant planes. To alleviate this issue, an adaptive label cost is used. The similarities between segmented planes are checked. If the angle and distance (parameter d in the plane equation) differences between segmented planes are less than given thresholds, n is replaced with the average number of inliers of the similar planes and their corresponding label costs are updated. Notably, very similar planes can be merged even under a small label cost.

2.3. Graph cuts based minimization

The above energy minimization problem can be resolved via graph cuts. The graph cuts technique constructs a graph for the energy function such that the minimum cut on the graph corresponds to the minimum of energy (Boykov et al., 2001; Kolmogorov and Zabin, 2004). We adopt a popular graph cuts based method, i.e., the extended α-expansion algorithm (Delong et al., 2012) to minimize the energy function in Eq. (1). Given a set of initial labels, the algorithm converts the multiple labeling problem into sequences of independent binary labeling problems and minimizes the energy function using graph cuts. Since the input to the energy function in Eq. (1) is a set of labels (planes), the planes derived from existing model fitting methods can be taken as the initial labels. In the paper, the plane fitting approach described in (Chauve et al., 2010) is adopted to produce initial planes without constraints on the plane normals. Algorithm 1 below illustrates the process of multi-plane optimization. Having initialized the three cost terms (steps 1–4), the graph cuts based method, i.e., the extended α-expansion algorithm is used to minimize the energy function (step 5) to find the optimal labeling of data points. Similiar to the PEARL (Propose Expand and Re-estimate Labels) algorithm (Isack and Boykov, 2012), the optimization is conducted in an iterative manner to refine the model paramters of the derived labels. Notably, some coplanar segments may be merged to a same label in the optimization. To separate them, a connectivity analysis needs to be performed after the label optimization (step 7).

Algorithm 1. Process of multi-label optimization.

Input: A set of lidar points P and its corresponding labeling L^0.
Input: Neighborhood system N.
Input: Threshold n and Δd.
Output: Optimal labeling L.
1. Initialize smooth costs (Eq. (4)) and set $t = 0$.
2. Derive the initial planes M_t from L^t.
3. Initialize data costs (Eq. (2)) and label costs (Eq. (5)).
4. Check the similarities between segmented planes and update their corresponding label costs.
5. Set $t = t + 1$ and run the extended α-expansion to compute the optimal labeling L^t.
6. If the energy decreases, go back to Step 2.
7. Compute the connected components of L^t and remove the labels with less than n points.
8. Update L^t and set $L = L^t$.

3. Experiments and results

Lidar data sets over Indianapolis (USA) and Vaihingen (Germany) are used for evaluation. The properties of the test data sets are summarized in Table 1. The test site in Indianapolis (Fig. 2a) is situated in downtown Indianapolis. It is characterized by complex pitched roofs consisting of small planes; some flat roof buildings are also present. The test site in Vaihingen (Fig. 2b) is located in the center of the city Vaihingen and characterized by

Table 1
Properties of the two test data sets.

Site	Indianapolis	Vaihingen
# Total points	87,862	151,177
#Roof points used	30,498	22,799
Equipment	Optech Gemini	Leica ALS 50
Acquisition date	December 2009	August 22, 2008
Coverage area	~39,840 m^2	~22,835 m^2
Point density	~1.4 pts/m^2	~4 pts/m^2
# Returns	2	3
# Returns used	2	3

#Returns: number of returns in lidar data; #Returns used: the number of used returns for segmentation; #Roof points used: the number of roof lidar points used for segmentation.

dense buildings with complex roof structures (Cramer, 2010). It is one of the benchmark data sets of the "ISPRS Test Project on Urban Classification and 3D Building Reconstruction" (http://www2.isprs.org/commissions/-comm3/wg4/tests.html), which aims to test various algorithms on urban object classification and building reconstruction.

We first use a sample building (Fig. 3) from the Vaihingen data set to illustrate the process of the label optimization. We add a number of random noise points to the raw building lidar points to evaluate the robustness of the algorithm. Fig. 3b presents the intial segmentation result, where many roof planes, including small ones are correctly separated from the noisy lidar points. However, spurious planes (e.g., segment C) are introduced and roof planes (e.g., segment A) may be partitioned into several segments. Similarly, segmentation of several transion regions between roof planes results in dangling portions (e.g., segment B). After the first iteration (Fig. 3c), most of the spurious planes are eliminated (cf. segment C in Fig. 3b) and the redundant planar segments belonging to the same roof plane are merged (cf. segment A in Fig. 3b). More noteably, the introduction of the smooth cost term can remove most of the dangling portions of planar segments (cf. segment B in Fig. 3b). In the subsequent iterations, the remaining arctifacts are gradually eliminated and the iteration converges after six times (Fig. 3e). The use of the label cost term may produce fewer labels, i.e., some coplanar roof planes are merged (e.g., planar segments D and E in Fig. 3e). Therefore, a connected component analysis of the segmented planes needs to follow to separate the coplanar segments. Given a distance threshold based on the point density, the components far away from other components are separated and assigned a new label. At the end, the labels (planes) consisting of less than a given number of lidar points are disregarded. Fig. 3f gives the final segmentation of the complex building. Fig. 4 shows the changes of the energy and the number of planes during the iteration. After the first iteration, most of the roof segmentation is completed since the number of identified planes and the energy drop sharply. In the subsequent iterations, only a few lidar points change their labels, leading to a quick convergence.

Same as the classical RANSAC, two necessary thresholds Δd (distance threshold) and n (the minimum number of lidar points required for a valid plane) must be selected before segmentation, where n is related to the density of the lidar points and the size of considered roof planes. In the paper, the distance thresholds in the initial segmentation and label optimization are set to be the same: $\Delta d = 7$ cm and 5 cm for Indianapolis and Vaihingen, respectively, and $n = 4$ for both data sets. Fig. 5 provides an overview of the segmentation results using the proposed approach. The Indianapolis data set (Fig. 5a) has 338 actual (reference) roof planes. Our approach misses 35 and yields a total of 337 planar segments, which includes 3 spurious planes and some over-segmented actual roof planes (e.g., roof planes A and B). The Vaihingen data set (Fig. 5b) has 241 actual (reference) roof planes. Our approach misses 10 and produces a total of 246 planar segments, including 5 spurious planes and some over segmented actual roof planes.

4. Assessment and discussion

Six metrics are used to evaluate the performance of the proposed approach. Completeness (Comp) is the percentage of reference roof planes that are correctly segmented. It is sensitive to the number of missed roof planes. Correctness (Corr) is the percentage of correctly segmented planes in the segmentation results. It is sensitive to the number of spurious planes. The two metrics are defined as

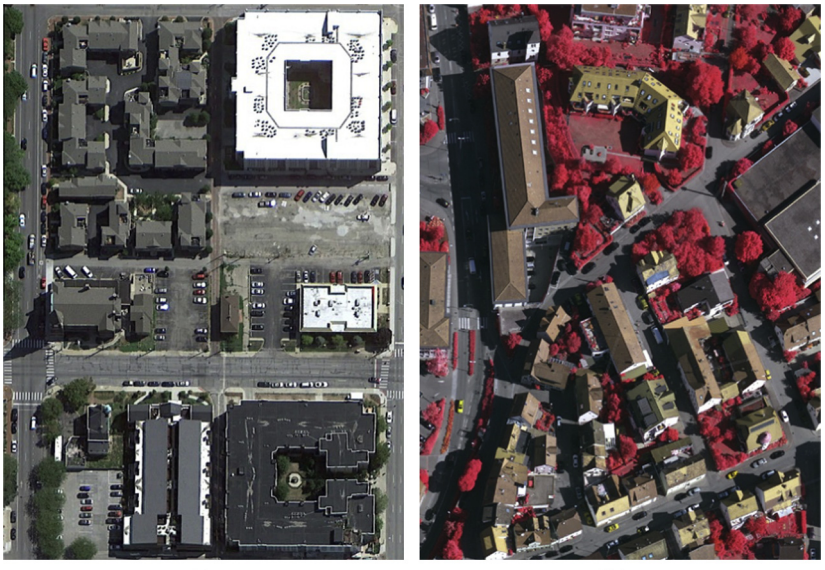


Fig. 2. Images of the two test sites (not in scale), (a) Indianapolis (from Google Map) and (b) Vaihingen (Cramer, 2010).

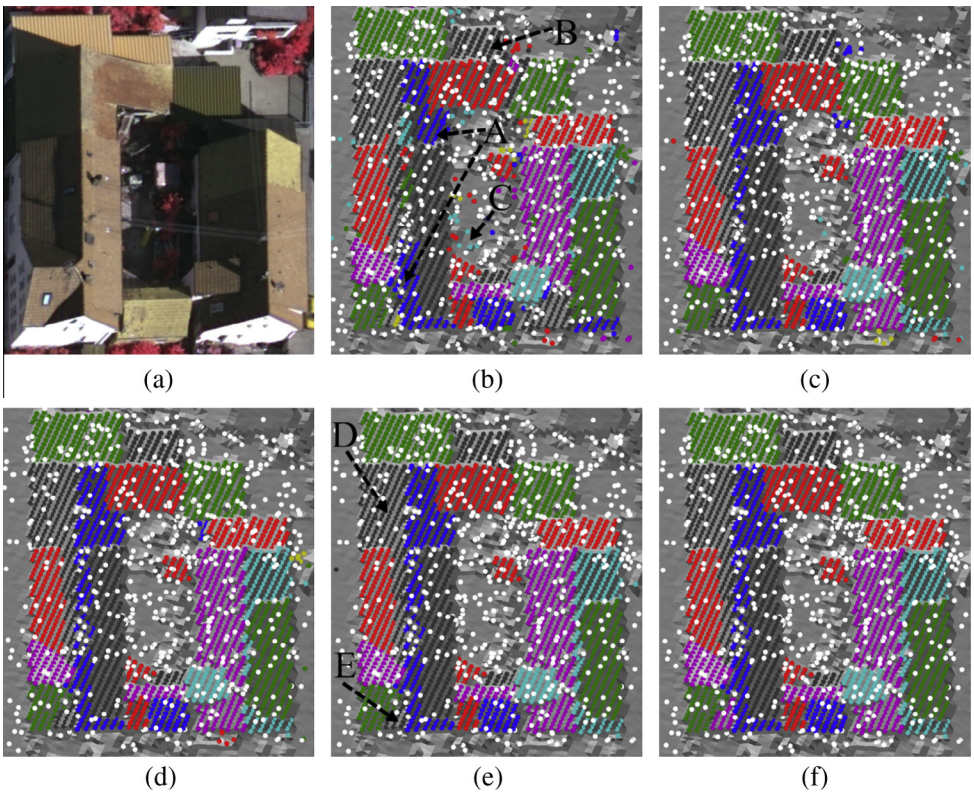


Fig. 3. Segmentation of roof lidar points (random noise accounts for 25% of the data): (a) reference image, (b) initial segmentation, (c) iteration 1, (d) iteration 3, (e) iteration 6 (convergence), and (f) final segmentation. Building point clouds are colored by segmented planar segments, with white dots for unsegmented lidar points.

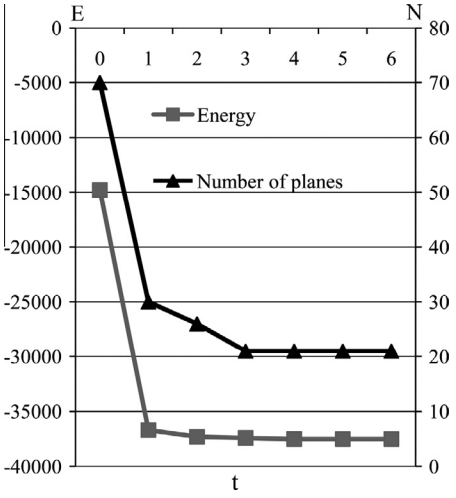


Fig. 4. Energy (E) and the number of planes (N) vs the number of iterations (t) in multi-label optimization.

$$comp = \frac{TP}{TP+FN}$$
$$corr = \frac{TP}{TP+FP} \quad (6)$$

where TP is the number of true positives, i.e., the number of planes found both in the reference and segmentation, FN is the number of false negatives, i.e., the number of reference planes not found in segmentation, and FP is the number of false positives, i.e., the number of detected planes not found in the reference. To be a true positive, a minimum overlap of 50% with the reference is required. To evaluate the effect of the incorrect segmentation, two additional metrics, detection cross-lap (DCL) rate and reference cross-lap (RCL) rate (Shan and Lee, 2005; Awrangjeb et al., 2010) are adopted. Detection cross-lap rate is the percentage of detected planes that overlap multiple reference roof planes. Reference cross-lap rate is the percentage of reference planes that overlap multiple detected planes. They are defined as

$$\mathrm{DCL} = \frac{N'_d}{N_d}$$
$$\mathrm{RCL} = \frac{N'_r}{N_r} \quad (7)$$

where N'_d is the number of detected planes that overlap more than one reference roof plane, N'_r is the number of reference roof planes that overlap more than one detected plane, N_r and N_d are respectively the number of reference planes and detected planes. However, the DCL and RCL metrics are mainly used for an object-based evaluation. As supplementary indicators, boundary precision and boundary recall are used (Estrada and Jepson, 2009). Boundary precision measures the portion of boundary points in segmentation that correspond to a boundary point in the reference. Boundary recall measures the portion of boundary points in the reference that correspond to a boundary point in the segmentation. For lidar points, the two metrics can be defined as

$$\text{Boundary precision} = \frac{|B_s \cap B_r|}{|B_s|}$$
$$\text{Boundary recall} = \frac{|B_s \cap B_r|}{|B_r|} \quad (8)$$

where B_r and B_s are respectively the set of boundary points in the reference and segmentation, $|\ |$ denotes the number of points in a data set. The reference planes and boundary points are derived by manually labeling (segmenting) the roof points.

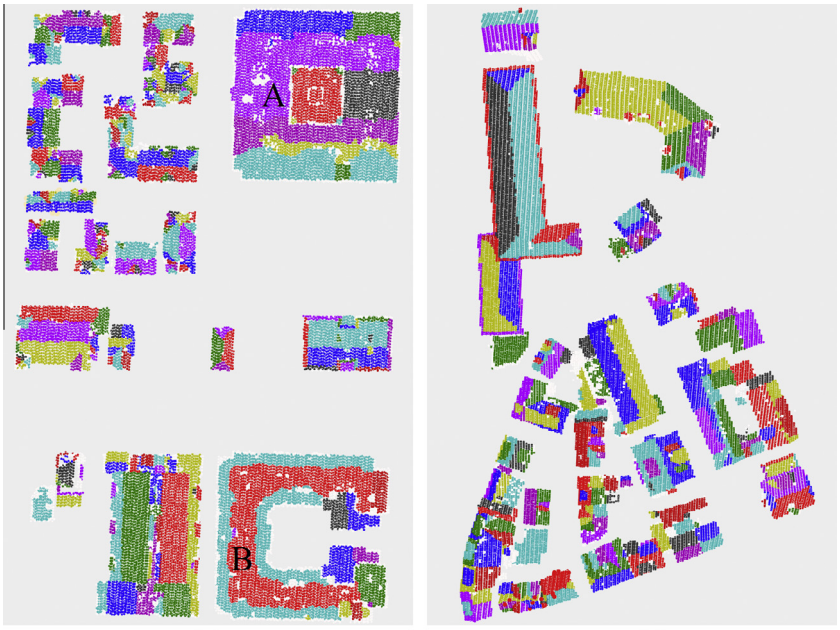


Fig. 5. Overview of the segmentation results color coded by planar segments: (a) Indianapolis and (b) Vaihingen.

4.1. Parameters sensitivity

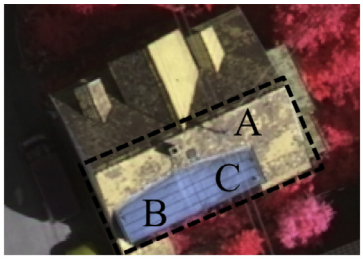
The geometric distance threshold (Δd) is an essential parameter in roof segmentation. Points within that distance threshold to a fitted plane become one segment. A large Δd results in coarse planes, whereas a small one produces fine segments. Incorrect segmentation tends to occur at plane transitions. Fig. 6 provides an example of segmentation under different distance thresholds. The angles between roof planes A and B, A and C, and B and C (cf. Fig. 6a) are respectively 34.0, 34.4, and 7.5 degrees. When the distance threshold is set to 7 cm and 15 cm, respectively, neither of the two initial segmentations (Fig. 6b and c left) can produce a satisfactory separation between plane B and C. However, the label optimization approach can yield a stable and correct boundary (Fig. 6b and c right) under both thresholds. What should be noted is the case of Fig. 6d where $\Delta d = 20$ cm. Such a large threshold makes plane B and C being segmented into one plane. The label optimization process does not contribute to this case since it cannot produce more number of planes than the initial segmentation can. Due to the lack of initial models corresponding to plane B and C, lidar points of these planes will be assigned to other similar planes after the label optimization.

4.2. Metric quality

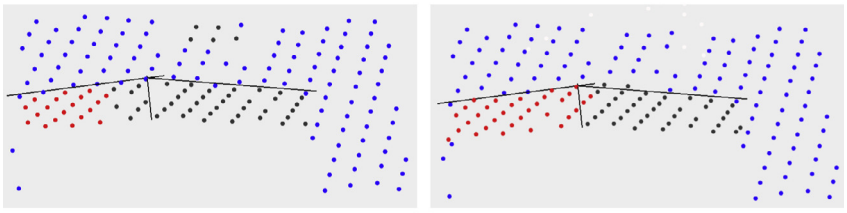
This section is primarily focused on the evaluation of completeness and correctness. Table 2 lists the statistics from the initial segmentation and its optimization. A number of observations can be noted based on this table. First, since the outcome of the initial segmentation defines the maximum number of plane segments (i.e. the set of possible labels), the optimization process cannot generate more number of plane segments. Under this restriction, the optimization process essentially may e.g., merge initial segments, split a segment and in the same time combine the results with other existing segments, or include initially not segmented lidar points to an existing segment, all under the restriction that no more number of segments is created. This fact is shown in the two #SP columns of Table 2 where less number of plane segments are present in the optimization outcome. Second, the optimization can segment many initially unsegmented lidar points into planar segments (#US columns). The remaining unsegmented points are 1.2–3.3 times less than the initial results. In average more than half (148 vs 64) initially unsegmented points (#US) are involved in the definition of the roof planes, which considerably improves the completeness from 80.1% to 92.3%. This is beneficial since more individual lidar points can contribute to the subsequent roof reconstruction. It should be noted that such property is likely due to the use of the smooth cost term (Eq. (1)) which enforces a transition at regions with sparse lidar data. Thirdly, many spurious planes resultant from the initial segmentation are resolved. Table 2 shows that the proposed optimization approach keeps the high completeness rate (92.3%) and in the same time improves the correctness from 93.0% to 100.0%, which means a very small false alarm rate can be expected.

Fig. 7 presents some representative segmentation results before and after the label optimization. As a common shortcoming of data-driven approach, the proposed approach fails to catch small roof planes with 0–6 lidar points (e.g., the elliptical regions in Fig. 7a, b, c and f). Similarly, thin planar faces with a width less than the point spacing are also missed due to insufficient lidar points. However, if the missed small roof planes are coplanar with, or very similar to a segmented one, they can likely be detected after label optimization. An example is given in Fig. 7c. Although the roof plane A (in the reference image) is missed in the initial segmentation result, it is correctly segmented through the subsequent label optimization. Due to the introduction of the adaptive label cost (Eq. (5)), the proposed approach behaves robustly even for complex buildings with a number of small roof planes. As shown in Fig. 7c, the test building consists of 49 roof planes, and a considerable portion of them has no more than 10 lidar points. Despite the existence of many complex shapes, 47 roof planes are correctly segmented from the lidar points, which improves the completeness from 77.6% to 95.9%.

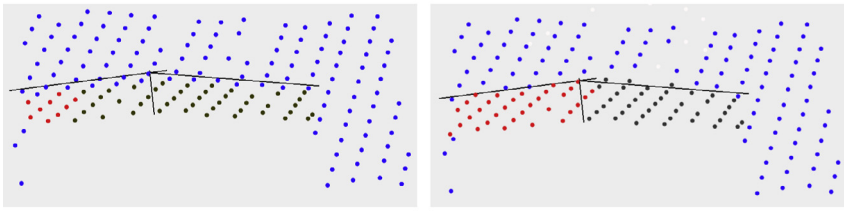
The label cost term encourages representing data compactly. Redundant planar segments will be merged or eliminated after the label optimization. Benefiting from the label cost, most of the spurious and over-segmented planes in Fig. 7 are eliminated or merged while retaining the small roof planes (e.g., the roof planes A and B in Fig. 7f). However, due to the presence of non-planar



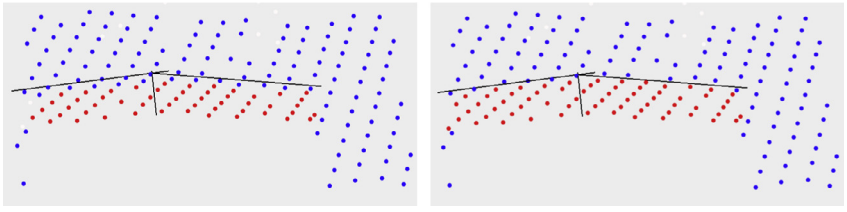
(a, reference image)



(b, Δd = 7cm)



(c, Δd = 15cm)



(d, Δd = 20cm)

Fig. 6. Segmentation under different distance thresholds (Δd) in plane fitting. From left to right: the initial segmentation and optimized segmentation. The three line segments represent the intersections of adjacent roof planes. Point clouds are colored by the segmented planar segments.

Table 2
Quality of segmentation results.*

Bld ID	#Pts	#Pls	Methods	#SP	#US	%Comp	%Corr	%RCL	%DCL	%BP	%BR
a	454	14	Initial	13	79	71.4	92.3	9.1	0	93.1	83.2
			Optimized	12	40	78.6	100	9.1	0	96.4	98.8
b	541	22	Initial	26	91	77.2	84.6	35.0	23.1	82.2	74.8
			Optimized	20	34	86.4	100	5.0	10.0	99.2	91.4
c	1161	49	Initial	58	252	77.6	93.1	25.5	22.4	81.0	65.2
			Optimized	47	79	95.9	100	6.4	0	99.8	97.2
d	1402	20	Initial	31	141	80.0	100	25.0	16.1	69.7	71.9
			Optimized	21	58	100	100	5.0	0	87.4	93.8
e	5180	13	Initial	27	251	92.3	85.2	7.7	11.1	82.0	79.0
			Optimized	12	75	100	100	0	7.7	97.4	96.0
f	425	12	Initial	14	86	75.0	100	33.3	7.1	93.4	63.7
			Optimized	11	28	91.7	100	0	0	97.3	91.6
g	633	12	Initial	20	73	76.9	95	30.8	5.0	90.1	73.4
			Optimized	16	25	92.3	100	7.7	0	95.3	91.9
h	2373	30	Initial	49	208	90.0	93.9	30.0	10.2	78.5	85.4
			Optimized	33	171	93.3	100	13.3	0	92.3	90.9
Average			Initial	–	148	80.1	93.0	24.6	11.9	83.8	74.6
Average			Optimized	–	64	92.3	100	5.8	2.2	95.6	93.9

* ID: building label (cf. Fig. 7); # Pts: the number of lidar points in a building; # Pls: the number roof planes in lidar data; Initial: initial segmentation results; Optimized: label optimization results; # SP: the number of segmented planes; #US: the number of unsegmented lidar points; %Comp: completeness; %Corr: correctness; %DCL: detection cross-lap rate; %RCL: reference cross-lap rate; %BP: boundary precision; %BR: boundary recall.

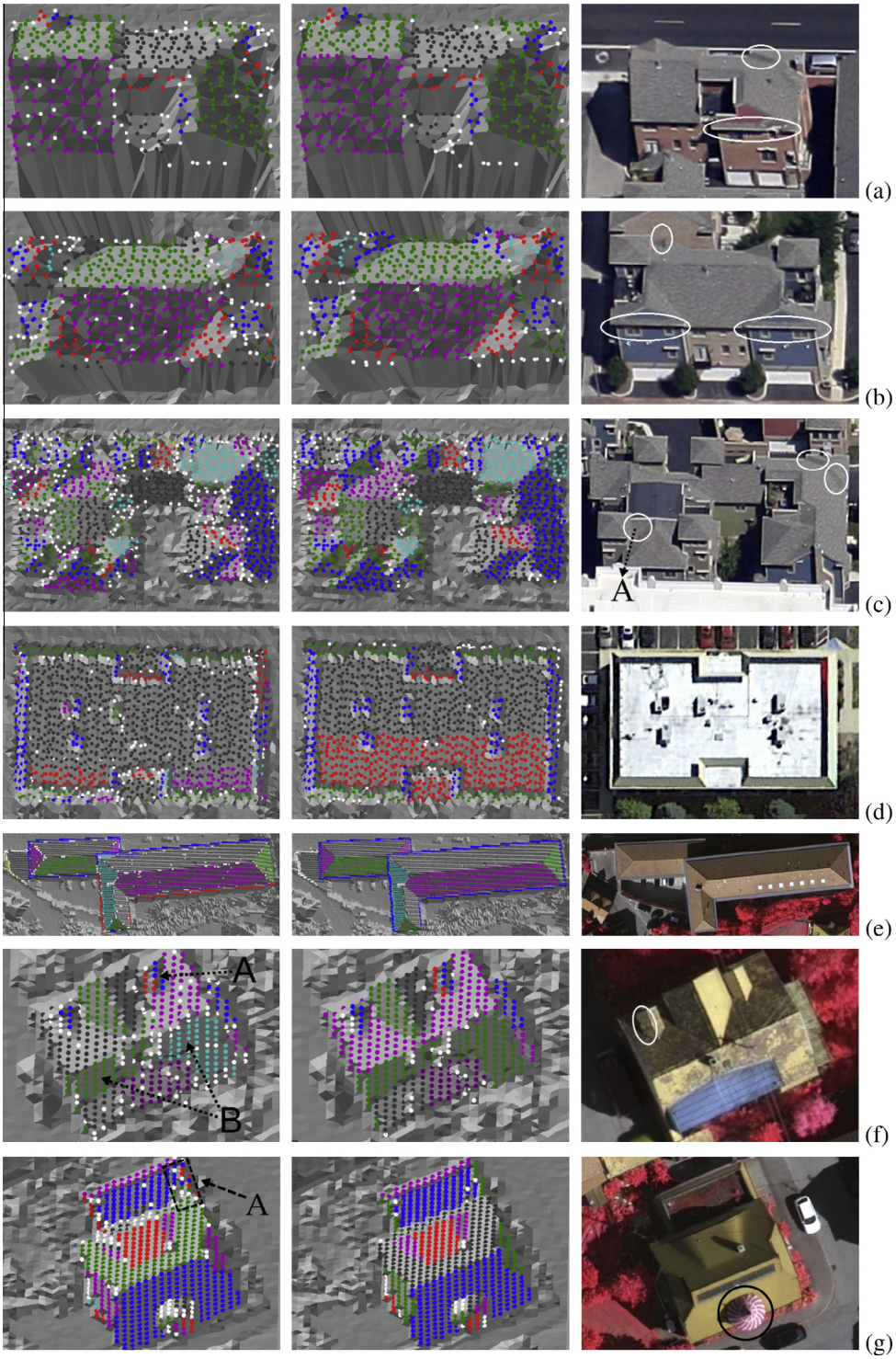


Fig. 7. Segmentation of roof lidar points. Building point clouds are colored by segmented planar segments, with white dots representing unsegmented points. From left to right: initial segmentation, the optimized results, and reference images. (a)–(d): Indianapolis (images from Google Map); (e)–(h): Vaihingen.

faces, the label optimization may fail to optimize the over-segmented planes derived from curved surfaces (cf. the cone roof in the reference image in Fig. 7g). Besides, since the distance threshold in the experiments may be small for some roof planes, there still exist over-segmented planes in the segmentation results. This problem is common for buildings with large flat roofs (cf. Fig. 7d and the roof planes A and B in Fig. 5a) (Oude Elberink, 2010). Additionally, the surface variations near roof eaves (cf. the regions B, C and D in Fig. 7h) may also result in over-segmentation.

4.3. Topologic quality

This section is primarily about detection cross-lap, where a detected plane overlaps multiple reference planes; and reference cross-lap, where a reference plane overlaps multiple detected planes. Conventional plane fitting methods usually extract planes one after the other from lidar points. Although these methods are robust to noise and outliers, they tend to result in mistakes at plane transitions, which in turn causes topological inconsistency

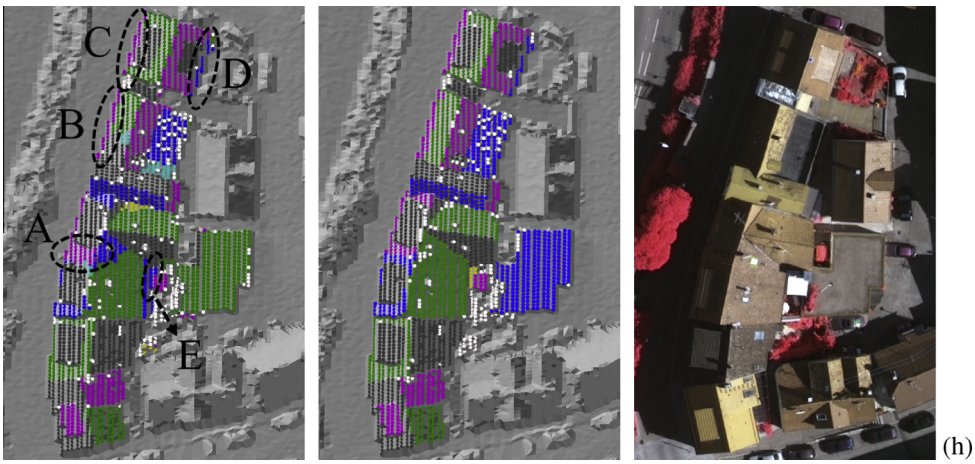


Fig. 7 (*continued*)

among segmented planes. Fig. 8 illustrates several examples of incorrect segmentation. Fig. 8a shows that two planes are merged into one. The type of incorrect segmentation in Fig. 8b is common (cf. the region A in Fig. 7h), where two planes are separated at a wrongly defined boundary. Fig. 8c is a typical over-segmentation at plane transitions (cf. segment E in Fig. 9d). Both Fig. 8b and c can be satisfactorily resolved through the label optimization process, however, the optimization of Fig. 8a may fail (cf. Fig. 6d) since the label optimization cannot increase the number of segments.

The topologic consistency can be measured by the detection cross-lap (%DCL) and reference cross-lap (%RCL) in Table 2. A good segmentation should have both as low as possible, so that the detected planes match the actual roof planes well and the segment boundaries adhere to roof edges. It is seen that the improvement is rather substantial since more than half of the detection cross-laps diminish after label optimization, whereas the improvement in reference cross-lap is significant, with an average reduction from 24.6% to 5.8%. Accordingly, all the boundary recalls (%BR) are increased by 15–32%, whereas the average boundary precision (%BP) is increased from 83.8% to 95.6%. A common topologic inconsistency in plane fitting is that the derived patches may cross other patches (Chen et al., 2012), which leads to dangling segments (i.e., broken boundaries) and results in a high detection cross-lap and reference cross-lap rates. As shown in Fig. 7g (left image), three planar segments cross each other at region A and the initial DCL, RCL, BP and BR (Table 2) are respectively 5%, 30.8%, 90.1% and 73.4%. After the label optimization, the topological consistency among the three planes is correctly retained and a low DCL and RCL rate (0% and 7.7%) and high boundary precision and recall (95.3% and 91.9%) are achieved in Table 2. Notably, such improvement is mainly due to the use of the smooth cost in optimization. Since neighboring points are encouraged to fit the same plane, the label optimization process can produce segments with compact boundaries, which in turn reduces the cross-lap rates among derived patches.

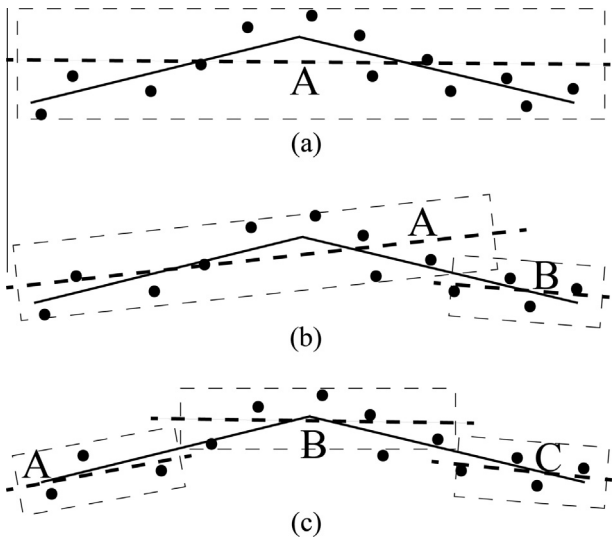


Fig. 8. Incorrect segmentations at a roof transition. Dots are roof lidar points; solid lines are roof planes; dash lines are fitted planes (dots in dashed rectangles are the inliers of fitted planes). (a) Adjacent roof planes are segmented into one segment. (b) A roof plane and its neighboring points of another plane are segmented into one segment. (c) The transition region is segmented into one segment.

The contribution of the label optimization to topologic consistency can be further demonstrated through Fig. 9. Fig. 9a and b are an overview of a selected area in Fig. 7h, respectively for its initial segmentation and label optimization. Fig. 9c–e are the blow-up views of the two segmentation results. As shown in Fig. 9d, most of the transition points are initially assigned to segment E, which leads to over-segmented planes H (Fig. 9c) and C. Meanwhile, some lidar points of segment F are wrongly assigned to segment C, yielding its incorrect dangling. Additionally, a spurious plane D consisting of lidar points near a roof ridge is also wrongly formed. Such topological inconsistencies can be resolved via label optimization. As shown in Fig. 9b, the most noticeable improvement is that the lidar points of plane E are correctly assigned to other segments such that the topological relationships among segments C, F, G and H are restored. Moreover, due to the use of the smooth cost, the dangling portions of segments A, B and C are removed, yielding a correct topological consistency.

The above discussion reveals that the proposed approach can produce spatially coherent segments and eliminate the dangling portions of segmented planes. Meanwhile, the number of spurious and over-segmented planes is reduced. Notably, in addition to the aforementioned artifacts, other factors such as roof matericals or features may also affect the segmentation quality. For example, some roofs may have a height texture of several centimeters (Oude Elberink and Vosselman, 2011), which may influence the selection of distance threshold in model fitting. Moreover, due to the lackness of sufficient returns, most of the returns from chimmey may be classfied into noise or outliers, whereas some skylights that are several centimetres above roof are segmented into roof planes.

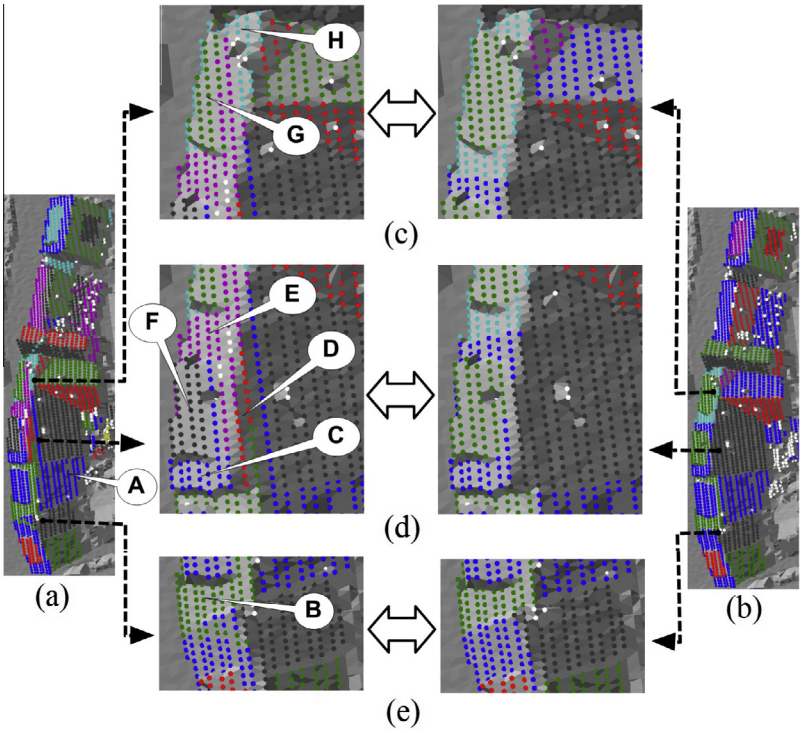


Fig. 9. Topologic consistency: (a) Initial segmentation (Δd = 10 cm). (b) The label optimization. (c–e) Comparisons of planar segments. The labeled letters represent the segmented planar segments. Building point clouds are colored by the segmented planar segments; white dots are unsegmented points.

5. Conclusions

This paper presents a global solution for roof segmentation from airborne lidar point clouds. Staring with an initial segmentation, the multi-label optimization is applied to optimize segmented planes. Experimental results from two test data sets show that 92.5% of the roof planes are correctly segmented, even for complex buildings. To be more specific, the completeness is increased from 80.1% to 92.3%, the detection cross-lap rate and reference cross-lap rate are reduced from 11.9% to 2.2%, and 24.6% to 5.8%, respectively. As a result, the noticeable incorrect segmentation problems at plane transition regions can be satisfactorily resolved and the topological consistency among segmented planes is correctly restored.

Compared to the existing model fitting methods for roof segmentation, the label optimization solution is a global one. Unlike the region growing and RANSAC-based approaches, the proposed approach can simultaneously determine multiple roof planes. A benefit of such approach is that multiple roof planes can compete for lidar points at their transition regions to reach an optimal segmentation. Another distinctive property of the proposed approach is that the spatial smoothness between data points is considered. It not only reduces the number of unsegmented points but also eliminates the incorrect dangling portions of the planar segments.

Notably, the proposed approach is a general solution to the roof segmentation task. To facilitate the evaluation and discussion, a plane fitting based on region growing (i.e., the initial segmentation) is used as a reference in this study, to which the proposed label optimization approach is compared. However, from the practical point of view, the label optimization method can also be used as a standalone method, which, as many clustering based segmentation methods, needs initial segmentation to start with. Therefore, it is expected that the proposed method can be combined with other existing ones with limited modification and significant performance enhancement.

However, the proposed solution also has limitations. It cannot produce more number of planes than what it starts with. This suggests an intentionally slight over-segmentation may be applied to start this method. Like many model fitting methods, a relatively small distance threshold would help for this end. Since the multi-label optimization solution provides a scalable framework for lidar point segmentation, additional constrains can be integrated into this formulation. For example, surface normal may be introduced to further consider the transition of planes and even non-polyhedral roofs.

Acknowledgement

The authors would like to thank the anonymous reviewers for their helpful comments and suggestions. This paper was partially supported by the National Basic Research Development Program of China (No. 2012CB719904).The Indianapolis data was downloaded from the 'IndianaMap Framework Data' (http://www.indianamap.org). The Vaihingen data was provided by the German Society for Photogrammetry, Remote Sensing and Geoinformation (DGPF [Cramer, 2010]: http://www.ifp.uni-stuttgart.de/dgpf/DKEP-Allg.html (in German).

References

Awrangjeb, M., Ravanbakhsh, M., Fraser, C.S., 2010. Automatic detection of residential buildings using LIDAR data and multispectral Imagery. ISPRS J. Photogram. Rem. Sens. 65 (5), 457–467.

Awwad, T.M., Zhu, Q., Du, Z., Zhang, Y., 2010. An improved segmentation approach for planar surfaces from unstructured 3D point clouds. Photogram. Rec. 25 (129), 5–23.

Boykov, Y., Veksler, O., Zabih, R., 2001. Fast approximate energy minimization via graph cuts. IEEE Trans. Pattern Anal. Mach. Intell. 23 (11), 1222–1239.

Bretar, F., Roux, M., 2005. Extraction of 3D planar primitives from raw airborne laser data: a normal driven RANSAC approach. In: Proceedings of IAPR Conference on Machine Vision Applications, Tsukuba, Japan, pp. 452–455.

Chauve, A.L., Labatut, P., Pons, J.P., 2010. Robust piecewise-planar 3D reconstruction and completion from large-scale unstructured point data, In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, pp. 1261–1268.
Chen, D., Zhang, L., Li, J., Liu, R., 2012. Urban building roof segmentation from airborne lidar point clouds. Int. J. Remote Sens. 33 (20), 6497–6515.
Cramer, M., 2010. The DGPF-test on Digital Airborne Camera Evaluation Overview and Test Design. Photogram. Fernerkundung, Geoinform. 2, 73–82.
Delong, A., Osokin, A., Isack, H.N., Boykov, Y., 2012. Fast approximate energy minimization with label costs. Int. J. Comput. Vision 96 (1), 1–27.
Dorninger, P., Pfeifer, N., 2008. A comprehensive automated 3D approach for building extraction, reconstruction, and regularization from airborne laser scanning point clouds. Sensors 8 (11), 7323–7343.
Duda, R.O., Hart, P.E., 1972. Use of the Hough Transformation to detect lines and curves in pictures. Commun. ACM 15 (1), 11–15.
Estrada, F.J., Jepson, A.D., 2009. Benchmarking image segmentation algorithms. Int. J. Comput. Vision 85 (2), 167–181.
Filin, S., Pfeifer, N., 2006. Segmentation of airborne laser scanning data using a slope adaptive neighborhood. ISPRS J. Photogram. Rem. Sens. 60 (2), 71–80.
Fischler, M.A., Bolles, R.C., 1981. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. Commun. ACM 24 (6), 381–395.
Gorte, B., 2002. Segmentation of TIN-structured surface models. Int. Arch. Photogram. Rem. Sens. Spatial Inform. Sci. 34 (4), 465–469.
Huang, H., Brenner, C., Sester, M., 2013. A generative statistical approach to automatic 3D building roof reconstruction from laser scanning data. ISPRS J Photogram. Remote Sens. 79, 29–43.
Isack, H., Boykov, Y., 2012. Energy-based geometric multi-model fitting. Int. J. Comput. Vision 97 (2), 123–147.
Kim, K.H., Shan, J., 2011. Building roof modeling from airborne laser scanning data based on level set approach. ISPRS J. Photogram. Remote Sens. 66 (4), 484–497.
Kolmogorov, V., Zabin, R., 2004. What energy functions can be minimized via graph cuts? IEEE Trans. Pattern Anal. Mach. Intell. 26 (2), 147–159.
Lari, Z., Habib, A., Kwak, E., 2011. An adaptive approach for segmentation of 3D laser point cloud. ISPRS – International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Calgary, Canada, pp. 103–108.
Melzer, T., 2007. Non-parametric segmentation of ALS point clouds using mean shift. J. Appl. Geodesy 1 (3), 159–170.
Oude Elberink, S.J., 2010. Acquisition of 3D topography: automated 3D road and building reconstruction using airborne laser scanner data and topographic maps. Ph.D thesis, University of Twente.
Oude Elberink, S.J., Vosselman, G., 2009. Building reconstruction by target based graph matching on incomplete laser data: analysis and limitations. Sensors 9 (8), 6101–6118.
Oude Elberink, S., Vosselman, G., 2011. Quality analysis on 3D building models reconstructed from airborne laser scanning data. ISPRS J. Photogram. Rem. Sens. 66 (2), 157–165.
Sampath, A., Shan, J., 2010. Segmentation and reconstruction of polyhedral building roofs from aerial lidar point clouds. IEEE Trans. Geosci. Remote Sens. 48 (3), 1554–1567.
Schnabel, R., Wahl, R., Klein, R., 2007. Efficient RANSAC for point-cloud shape detection. Comput. Graph. Forum. Wiley Online Library, 214–226.
Shan, J., Lee, S.D., 2005. Quality of building extraction from IKONOS imagery. J. Survey. Eng. 131 (1), 27–32.
Tarsha-Kurdi, F., Landes, T., Grussenmeyer, P., Koehl, M., 2007a. Model-driven and data-driven approaches using lidar data: analysis and comparison. International Archives of Photogrammetry, Remote Sensing and Spatial Information Systems, 87–92.
Tarsha-Kurdi, F., Landes, T., Grussenmeyer, P., 2007b. Hough-transform and extended ransac algorithms for automatic detection of 3d building roof planes from lidar data. Int. Arch. Photogram., Rem. Sens. Spatial Inform. Syst. 36, 407–412.
Verma, V., Kumar, R., Hsu, S., 2006. 3D building detection and modeling from aerial LIDAR data, IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, New York, NY, USA, pp. 2213–2220.
Vitti, A., 2012. The Mumford-Shah variational model for image segmentation: an overview of the theory, implementation and use. ISPRS J. Photogram. Rem. Sens. 69, 50–64.
Vosselman, G., Dijkman, S., 2001. 3D building model reconstruction from point clouds and ground plans. Int. Arch. Photogram. Remote Sens. Spatial, Inform. Sci. 34(3/W4), 37–44.
Yan, J., Jiang, W., Shan, J., 2012. Quality analysis on RANSAC-based roof facets extraction from airborne LIDAR data. ISPRS-Int. Arch. Photogram. Rem. Sens. Spatial Inform. Sci. 1, 367–372.
You, R.-J., Lin, B.-C., 2011. Building feature extraction from airborne lidar data based on tensor voting algorithm. Photogram. Eng. Rem. Sens. 77 (12), 1221–1232.
Zhang, K., Yan, J., Chen, S.-C., 2006. Automatic construction of building footprints from airborne LIDAR data. IEEE Trans. Geosci. Remote Sens. 44 (9), 2523–2533.