HYBRID IMAGE SEGMENTATION USING LIDAR 3D PLANAR PRIMITIVES

Frédéric Bretar [a], Michel Roux [b]
[a] Institut Géographique National
2-4 Av. Pasteur 94165 St. Mandé cedex, France
Email: Frederic.Bretar@ign.fr
[b] GET-Télécom Paris - URA 820 LTCI - Département TSI
46 Rue Barrault 75013 Paris, France
Email: Michel.Roux@enst.fr

Commission III, Working Group 3

KEY WORDS: Hierarchical segmentation, Lidar 3D data, watershed, RANSAC, Extended Gaussian Image, data fusion

ABSTRACT:

Airborne laser systems are nowadays well-known to provide regular and accurate altimetric data. The aim of this paper is to investigate the potential of using extracted features from a Lidar 3D point cloud for data fusion purpose through an hybrid image segmentation algorithm. The general context of this study is the building reconstruction. We first describe an efficient algorithm for extracting 3D planar primitives from a laser survey over urban areas. It is based on a normal driven random sample consensus (ND-RANSAC) which consists of randomly selecting sets of three points within laser points sharing the same orientation of normal vectors. A robust plane is then estimated with laser points that are likely to belong to the real roof facet. The number of random draws is managed automatically with a statistical analysis of the distribution of normal vectors within an approximation of the Gaussian sphere of the scene. These 3D facets are then introduced into an image segmentation algorithm based on a bottom-up region merging scheme. Initial regions are computed by a watershed transform onto the gradient of the aerial image. Regions adjacency are represented with a Region Adjacency Graph (RAG). Series of successive merging generates a hierarchy of RAGs where edges are mutual inclusion relationships. The process terminated when the entire image is represented as a single region. A cut in the hierarchy provides a desired image partition. Results of the facet extractions as well as of the segmentation are shown and discussed.

1 INTRODUCTION

Airborne laser technologies, with the generation of 3D point clouds, have been the crux of intensive researches for the last past years. Scientists did work onto various themes as the study of the final accuracy of the point cloud (Bretar et al., 2003), the strip adjustment problem (Filin, 2003), the automatic classification (Sithole and Vosselman, 2003), the extraction of characteristic features and more recently the fusion of laser data with external sources. A laser survey is classically composed of 3D points with an irregular spacial distribution (depending on the laser system) representing the topography of a certain landscape. It is a basic geometric representation of the scene where no key is provided to help its understanding. An intermediate object recognition level consists of grouping primitives (3D entities) originating from the same object (Rottensteiner and Briese, 2003). When studying a urban landscape, roof facets become extremely important when representing a 3D scene. More over, 3D facets are particularly relevant information for data fusion purposes. Is the laser-based geometry sufficient for high level landscape descriptions as full building reconstruction approach or realistic rendering? The detection of 3D planar patches within a point cloud generally provides sparse planes with gaps where the information is lost or at least not directly exploitable. Retrieving adjacency between this set of detected plane is not immediate neither. It seems that information provided by a radiometric description (aerial/satellital images) of the landscape through a photogrammetric analysis (oriented aerial image, automatic digital elevation model generation with correlation computations...) are complementary data sources for a full landscape description. Schenk (Schenk and Csatho, 2002) describes a possible combination of aerial imagery and LIDAR data. Planar surface patches are defined with a region-growing segmentation. Possible adjacent planes are intersected while image edges are detected with a Canny operator. Some plane intersection are confirmed by fusionning edge information. McIntosh (McIntosh and Krupnik, 2002) proposes to detect (with an optimal zero-crossing operator) and to match edges in aerial images to refine the digital surface model produced from airborne scanner data. Sohn (Sohn and Dowman, 2003) uses edges detected in an IKONOS image to constraint the search of a polyhedral building shape. The aim of this paper is to investigate the potential of a deep synergy involving 3D extracted facet from a LIDAR point cloud and aerial images throughout an hybrid image segmentation algorithm.

We would like to focus the present study onto two main points. On the first hand, in a building reconstruction approach, we consider the detection of 3D planar primitives over a set of *connected* buildings. This problem have been tackled by many authors. Maas and Vosselman (Maas and Vosselman, 1999) proposed a first solution involving invariant moments of point clouds. A second approach, also studied by Hofmann (Hofmann et al., 2003), consists of detecting planar faces in a triangulated point set, investigating the parameter space of planes. Plane directions are extracted through a 3D cluster analysis on the Gaussian sphere. An obvious disadvantage of this technique is that parallel planar faces cannot be separated directly on the Gaussian sphere. Pottman (Pottman et al., 2002) proposes to use a special distance to measure the proximity of planes and then to enhance the 3D clustering process. An

other approach (Vosselman and Dijkman, 2001) is based on the well-known 3D Hough transform to detect planar faces from the irregularly distributed point clouds. Finally, the region-growing segmentation based on planarity criteria (normal vectors, residuals of a local fitted plane,...) are wide-used (e.g. (Rottensteiner and Briese, 2003)). We propose an alternative solution to detect roof facets of buildings based on a normal driven RANSAC (RANdom SAmple Consensus) related approach which is described in section 2.

On the second hand, we propose to insert these primitives into an hybrid image segmentation algorithm based on a bottom up region merging scheme (Haris et al., 1998), (Ward, 1963). This approach is particularly well adapted to our data fusion problem since a cost function has to be defined including both radiometric and geometric (related to the 3D facets) information. We will see in section 3 that this segmentation is based on operations performed onto a Region Adjacency Graph (RAG). A hierarchy is built up as the region merging process goes along. A cut of the hierarchy at a defined level provides an appropriate segmentation. At last, some results show the potential of this approach. We will also discuss problematic points related to the cut's shape.

2 EXTRACTION OF 3D ROOF FACETS

2.1 Background

The RANSAC algorithm introduced by Fischler and Bolles (Fischler and Bolles, 1981) with applications to the context of roof facet detection would be formulated as follow: randomly select a set of N plans (3 points m) within a point cloud $\mathcal{S}$ and keep memory of the number of points (supports) which distance from the associated planes are less than a critical distance (d). A least square estimation of the final plane ($\mathcal{P}_{final}$) is performed with the set of supports ($\mathcal{M}_{final}$) belonging to the plane with the highest score. The set $\mathcal{M}_{final}$ is then extracted from the initial point cloud $\mathcal{S}$. The algorithm runs until $card(\mathcal{S}) < 3$.

Algorithm 1 Basic RANSAC for detecting roof facets

repeat
  while $n \leq N$ **do**
    Randomly select a plan $\mathcal{P}$ (3 points)
    $\mathcal{M} = \{m \in \mathcal{S} / ||m - \mathcal{P}(m)|| \leq d\}$
    $\mathcal{M}_{card(\mathcal{M})} \leftarrow \mathcal{M}$
    $n = n + 1$
  end while
  $\mathcal{M}_{final} = \arg\max_{n \in \mathbb{N}} \mathcal{M}_n$
  $\mathcal{P}_{final} = \arg\min_{\mathcal{P}'} \sum_{m \in \mathcal{M}_{final}} ||m - \mathcal{P}'(m)||^2$
  $\mathcal{S} \leftarrow \mathcal{S} \backslash \mathcal{M}_{final}$
until $card(\mathcal{S}) < 3$

This approach may be extremely time consuming since we must ensure a minimum number of draws (N) so that a correct plan $\mathcal{P}$ should be instantiated. Most often it is not worthy to try all possible draws (Hartley and Zisserman, 2000). In other words, for a given probability p of drawing a correct plane $\mathcal{P}$ (that is three points without outlier), we would like to maximize the probability w that any selected point is an inlier (w^3 for 3 points). p, w and N are related to each other by the following equation :

$$(1-p) = (1-w^3)^N \Leftrightarrow N = \frac{log(1-p)}{log(1-w^3)} \qquad (1)$$

N can therefore be calculated directly from the knowledge of p and w. p is generally kept constant to 0.99. The main idea of this plane detection algorithm is that searching for the roof facets where they really could be located should highly improve the efficiency of a blind RANSAC approach. In our context, main plane directions correspond to roof facet orientations. As a result, focusing the consensus onto regions sharing the same normal orientation will constrain the probability w to follow specific statistical rules as developed in section 2.2.

2.2 The ND-RANSAC algorithm

Point clustering based on surface normal estimation There are several methods for obtaining local surface normal from range data (McIvor and Valkenburg, 1997). These normals are then processed for segmenting planar surface regions of range images like the fast segmentation method of Taylor (Taylor et al., 1989). It is a split-and-merge method, where the homogeneity criterion is based on the comparison of two angles describing the normal orientation and the original range value. Merging is based on simple minimum and maximum value comparison of neighboring regions.

Here, we propose a planar segmentation of the laser point cloud by analyzing the Gaussian sphere (GS) of the scene. Normal vectors are calculated over a regular grid by extracting a circular neighboring of the central 3D point. A plane is then estimated using a robust regression of M-estimators' family with the norm $L_{1.2}$ (Xu and Zhang, 1996). The mass density of the normal vectors on the GS is described by an extended Gaussian image (EGI) (Horn, 1984). First, the GS can be approximated by a tessellation of the sphere based on regular polyhedrons. Such tessellation is computed from a geodesic dome based on the icosahedron divided into f sections (f is a power of 2). The EGI (figure 1) can be computed locally by counting the number of surface normals that belong to each cell. The values in the cells can be thought as an histogram of the orientations.

The angular spread (related to the number of faces) depends on the error we tolerate for the coherence of the normal vectors in a cell. As a result, within a specific cell, normals $\begin{pmatrix} n_x \\ n_y \\ n_z \end{pmatrix}$ will be distributed following a certain density of probability $p(n_x, n_y, n_z)$, which will be analyzed in the next section. Each cell with a minimum number of laser points is linked to the corresponding image of normal vectors. Regions sharing the same normal orientation are detected. They are then labeled providing a set of clusters which are ordered depending on their surface. Only most represented orientations (largest areas) will be treated by selecting laser points belonging to these areas. The RANSAC algorithm is a general robust approach to estimate models. Instead of using the whole data to obtain an initial solution and then attempting to eliminate the invalid data point, RANSAC uses a data set as small as feasible and enlarges this set with consistent data when possible. Two parameters have to be tuned: the critical distance and the number of draws. The first one depends on the noise ratio of the data while the second one depends on the statistical distribution of points onto the Gaussian sphere.

The critical distance We noticed in section 2.1 that supports were considered in the set $\mathcal{M}$ only if their distance to the associated random plane was less than a critical distance d. This distance may be seen as the standard deviation of the supports with regard to the 3D plane. d is therefore defined for each cluster $\mathcal{C}$ as proportional to the final residual square root of a least square fitted plane

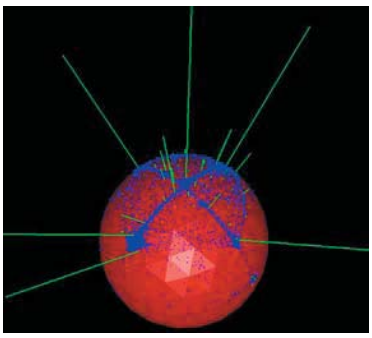


Figure 1: Extended Gaussian Sphere of the building represented in figure 3(b)

estimated from the entire laser points $\{P\}$ within $\mathcal{C}$. If $\{P'\}$ is the orthogonal projection of $\{P\}$ onto the fitted plane, then

$$d = \sqrt{\sum_{P \in \mathcal{C}} \|P - P'\|^2} \tag{2}$$

The number of draws The number of draws to be performed depends on the distribution $p(\overrightarrow{n})$ of the normal vectors within a cluster (where $\overrightarrow{n} = \begin{pmatrix} n_x \\ n_y \\ n_z \end{pmatrix}$ are considered as three random variables), and especially on the probability w that any selected point is an inlier (equation 1). Following the definition of a probability and considering that the final plane will be close to the mathematical expectation $\mathbf{E}(\overrightarrow{n})$ of the distribution, w satisfies:

$$w = \int_{\mathbf{E}(\overrightarrow{n})-\overrightarrow{\sigma}}^{\mathbf{E}(\overrightarrow{n})+\overrightarrow{\sigma}} p(\overrightarrow{n})\,\mathrm{d}\overrightarrow{n} \tag{3}$$

where $\overrightarrow{\sigma}$ is the standard deviation of the distribution. We may assume that the three random variables n_x, n_y, n_z are independent to write:

$$w = \int_{\mathbf{E}(n_x)-\sigma_{n_x}}^{\mathbf{E}(n_x)+\sigma_{n_x}} p_x(n_x)\,\mathrm{d}n_x . \int_{\mathbf{E}(n_y)-\sigma_{n_y}}^{\mathbf{E}(n_y)+\sigma_{n_y}} p_y(n_y)\,\mathrm{d}n_y . \int_{\mathbf{E}(n_z)-\sigma_{n_z}}^{\mathbf{E}(n_z)+\sigma_{n_z}} p_z(n_z)\,\mathrm{d}n_z . \tag{4}$$

$p_x(n_x)$ (resp. $p_y(n_y), p_z(n_z)$) is explicitly calculated as the derivative of the empirical probability density function F_{K_x} (resp. F_{K_y},F_{K_z}) with:

$$F_{K_x}(x) = \begin{cases} 0 & \text{if } x < \mathbf{inf}\, n_i \\ \frac{n_i}{K_x} & \text{if } n_i < x \leq n_{i+1} \\ 1 & \text{if } x > \mathbf{sup}\, n_i \end{cases} \tag{5}$$

where n_i is the proportion of values less than x and K_x (resp. K_y,K_z) the number of realizations of the random variable n_x (resp. n_y, n_z). The performances of this detector is discussed in section 5.

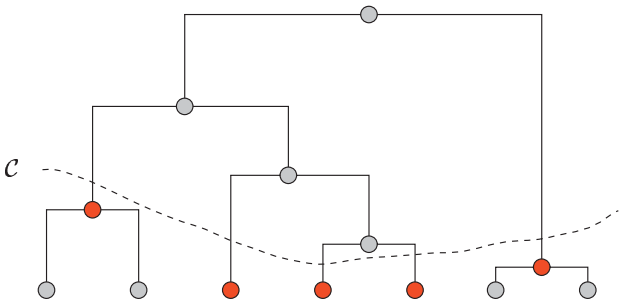


Figure 2: Tree structure of $\mathcal{H}$ represented as a dendrogram. The dot line corresponds to a given cut $\mathcal{C}$ composed of red nodes.

3 Image Segmentation

3.1 Hierarchical region merging scheme

Segmenting an image $\mathcal{I}$ consists of determining a partition $\Delta_N(\mathcal{I})$ of N regions $R_{i\in[1,N]}$ (a region is a connected set of pixels that satisfies certain predefined homogeneity criteria) satisfying:

$$\Delta_N(\mathcal{I}) = \bigcup_{i\in[1,N]} R_i\,,\; R_i \bigcap_{i\neq j} R_j = \emptyset,\; \forall i, R_i \text{is connected}$$

It is clear that a unique and relevant *a priori* partition of a given image does not exist. The interpretation of an image can be performed at a given scale of details for a specific purpose. Here, we would like to obtain a partition whereon roof facets are clearly delineated and understandable, that is interpreting the image at the roof facet scale. We therefore retained a region-based approach organized as a hierarchical description of the scene. A hierarchy over a set is a set of mutually included subsets. In our description, a given scale is reached after the fusion of two regions. The inclusion function (or the inter-scale relation) is the appearance level of a region pair. From an initial image partition represented by elementary regions, similar ones are successively merged according to certain decision rules. Such a segmentation paradigm can be represented by a tree structure (a hierarchy) where regions (the tree's nodes, cf. figure 2) at lower levels are joined to form regions at higher levels. Tree branches indicates region set inclusions. The root of the tree corresponds to image $\mathcal{I}$ and the leaves to the image initial segmentation. A hierarchy $\mathcal{H}$ can be considered as a union of partition sets. A *cut* of $\mathcal{H}$ provides a partition of $\mathcal{I}$. More formally, a *cut* of $\mathcal{H}$ is a set of nodes that intersects once and only once all branches of $\mathcal{H}$. A cut is therefore not necessarily horizontal.

A data structure for representing image partitions is the RAG. The RAG of a N-partition (a N-RAG) is defined as an undirected graph $G = (V, E)$ where V ($card(V) = N$) is the set of nodes and E the set of edges. Each region is represented by a graph node. An edge represents the adjacency of two regions. A score is affected to each edge. This score measures the "similarity" between two adjacent regions. The general idea of a hierarchical ascendant segmentation is to merge the most "similar" pair of regions at each iteration until a single region remains. Algorithm 2 gives the outlines of the hierarchical region merging scheme onto a RAG.

Different similarity functions can be employed, each one corresponding to different definitions of the picture segmentation task. Here, we want to detect roof facets within an aerial image what consists of partitioning the image with relevant planar patches. This

Algorithm 2 Hierarchical region merging onto a RAG

```
create the initial N-RAG
n = 0
while n < N do
    Find the minimum cost edge in the (N-n)-RAG
    Merge the corresponding pair of regions
    update the (N-n-1)-RAG edge attributes
    n = n + 1
end while
```

segmentation process involves both geometric (planes are not directly visible onto a single image) and radiometric features (color-like neighboring regions tend to belong to the same plane).

3.2 Region similarity function

We used the Ward criterion to manage with color-like neighboring regions. This model supposes a piecewise constant representation of the image. It consists of approximating each region R_k by its mean $\mu(R_k)$ with

$$\mu(R_k) = \frac{1}{||R_k||} \sum_{i=0}^{||R_k||} \mathcal{I}(p_{k,i})$$

where $||R||$ is the cardinal of the region and $p_{k,i}$ the position of the pixel i within the region R_k. We show (Haris et al., 1998) that if Δ_K^* is an optimal K-partition with respect to the total squared error $Err(\Delta_K^*)$ with

$$Err(\Delta_K^*) = \sum_{k=0}^{K} \sum_{i=0}^{||R_k||} \left(\mathcal{I}(p_{k,i}) - \mu(R_k)\right)^2$$

then the optimal $(K-1)$-partition is generated by merging the pair of regions of Δ_K^* which minimizes the function:

$$\delta(R_i, R_j) = \frac{||R_i||.||R_j||}{||R_i|| + ||R_j||} \left(\mu(R_i) - \mu(R_j)\right)^2$$

where R_i and R_j are neighboring regions.

The geometry of lidar data is integrated into adjacency relationships through the detected 3D facets (section 2). The set of 3D facets is denoted $\mathcal{F}$ in the following. Introducing different criteria into the edge cost function will guide the fusion process through the apparition level (scale) of region pair fusion. In this hierarchical segmentation scheme, we will favor at first the merging of neighboring regions which belong to the same 3D facet. From an initial segmentation, successive fusions will be operated under geometric constraints considering that 3D planar primitives detected into the laser point cloud are first order information. Once the geometric contribution is entirely exploited, regions will merge on the basis of radiometric considerations. We have mentioned that facet frontiers were not accurate enough since they are calculated from sparse 3D points without any knowledge onto facades nor ground plane. The geometric constraints will therefore be applied only to regions strictly included into $\mathcal{F}$ (equation 6). Regions belonging to two different facets are strongly penalized so that roof facets should merge at the top of the hierarchy (equation 7).

The final similarity function $\mathcal{S}$ can be expressed as follow:

$$\mathcal{S}(R_i, R_j) = \frac{||R_i||.||R_j||}{||R_i|| + ||R_j||} \Big(\left(\mu(R_i) - \mu(R_j)\right)^2 + \alpha\, \zeta(R_i)\zeta(R_j)\mathcal{B}(R_i, R_j) \Big)$$

where

$$\zeta(R_k) = \begin{cases} 1 & \text{if } \exists f \in \mathcal{F} / R_k \in f \text{ strictly} \\ 0 & \text{if not} \end{cases} \quad (6)$$

and

$$\mathcal{B}(R_i, R_j) = \begin{cases} -1 & \text{if } \exists f \in \mathcal{F} / R_i \cap R_j \in f \\ 1 & \text{if not} \end{cases} \quad (7)$$

α is a weighting factor. A high α will favor geometric merging at the beginning of the process and will postpone the fusion of image-based building roofs at its top-end. When a region R_k becomes large enough, $\zeta(R_k)$ is null and regions are exclusively merged onto radiometric similarities.

The initial segmentation An initial segmentation contains elementary regions to be merged. Parts of their boundaries define region boundaries of a particular partition. Objects are generally isolated by strong intensity jumps (image edges). We will constrain this initial segmentation to follow the image edges performing a watershed transform onto the gradient image (Canny-Deriche operator). The output of the watershed algorithm is a tessellation of the input image into its different basins, each one characterized by a unique label.

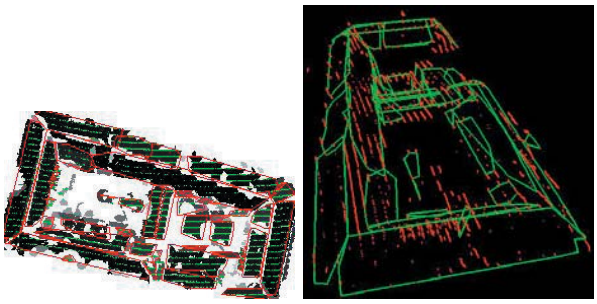
4 The Data Set

Laser data have been acquired over the city of Amiens, France, by the company TopoSys©in 2002. This firm owns a self-made lidar acquisition system, which is composed of two rigid blocks of optical fibers (emission and reception of laser pulses). The ground pattern of laser impacts is strongly irregularly distributed. The spatial density of the point cloud is roughly one point every $10\,cm$ along the flight track and one point every $1.2\,m$ in the cross-track direction. The ground density is 7.5 points/m^2. We worked onto a single strip which have been filtered into ground and non-ground points (Bretar et al., 2004). A simple morphological analysis allows to isolate sets of connected buildings.

For the present study, we did use aerial color images fully corrected of altimetric distortions (they are ortho-rectified) taken in 2001 over the same city of Amiens.

5 Results and Discussion

We will present on going results onto a complex building and discuss at first the results of the 3D facet detector. Figure 3(b) shows the 3D planar primitives calculated with the ND-RANSAC algorithm over a building. They have been computed with the following parameters: normals are calculated over a $0.3\,m$ resolution grid onto a circular neighborhood of $2\,m$ radius, the minimum number of laser points within a cluster is set to 30. Finally, each cluster must be composed of at least 25 pixels to be considered. This methodology depends mainly on the quality of the normal vector map and as a result of the spacial distribution of laser points. The irregularity of the ground pattern necessitates to integrate laser points onto a large neighborhood so that enough points should be considered



(a) (b)

Figure 3: (a) is the cluster images coded in gray level whereon laser points have been projected (green points) as well as retrieved facets (red polygons). (b) is the 3D representation of roof facets (green polygons). Red points are the residuals of laser points which have to be considered as support of any plane.

into the plane estimation process. Nevertheless, a large neighborhood tends to smooth the normal information, especially for instant at the building roof tops. The point density, which is a limitation factor, should be enhanced by crossing different strips or joining a raster representation of the topography (either from the laser itself or from a photogrammetric DSM). If our methodology is slightly less general than a classical RANSAC which provides, whatever happens, relevant planes within the point cloud after theoretically an infinite amount of iterations, it is much more time efficient. It is also particularly interesting to take benefit of the specific building geometry.

The approximation of the Gaussian sphere by a regular polyhedron, that is its discretization with a fixed step, has advantages with regard to other methods like for instant the K-means one. The main one is no doubt to manage with the extension of each cell which is the actual tolerance (the distribution extension) granted around one direction. A rough discretization provides large areas of similar normal vector. It is exactly what we are looking for. Indeed, this clustering stage is only considered as a focusing stage. Therefore, there is no need to have accurate boundaries, for the final facet estimation is performed directly onto the point cloud. As a result, the best parameters include a not-so-fine GS approximation to provide large focusing areas. Analyzing the distribution of the 3D points within these focused areas will provide the best robust plane. This approach will not provide all possible facets, but the detected ones are relevant.

The described approach for the data fusion of lidar and aerial images proposes to combine a real 2D description of images unlike other approaches based on the detection of lineic features, and full 3D laser data. It provides an interpretation of the image for building reconstruction purposes. More precisely, a region-based image description gives 2D adjacency relationships of facets which can directly be related to the 3D facet orientations.

Figure 4 shows the initial segmentation (red contours) superimposed onto the ortho-image to be segmented. The actual implementation is based on a watershed of the gradient image. It could be interesting as future work to combine other image sources, especially images of the same building taken from different view angles. Combining the gradient of such images may lead to an ortho image of gradients. It is calculated by accumulating the gradient image magnitudes in an orthogonal projection. Edges invisible in certain images may appear in others and finally make the watershed more relevant.

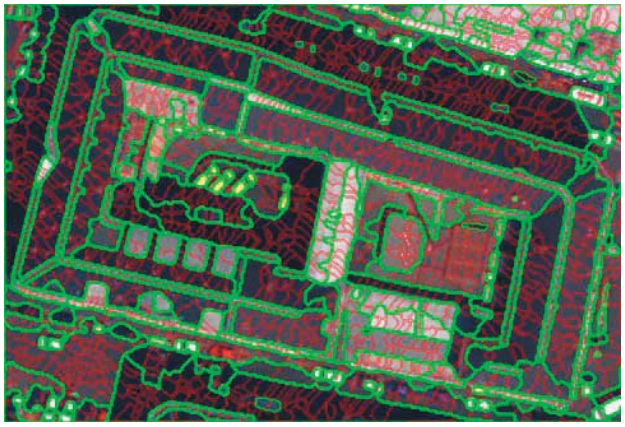


Figure 4: Segmented ortho-image where contours of watershed regions (red) of the corresponding gradient image are superimposed. A cut in the hierarchy (performed on visual criteria) represents the depicted image partition where region contours are drawn (green lines are the contoured regions of figure 5).

Figure 5 shows the label image corresponding to an horizontal cut of the hierarchy, also represented in figure 4 as green contours. Most of building facets are interpreted as large labeled areas and make the building structure clearly visible. We can remark that roof tops and gutters are not defined as a single line in the watershed, but as two parallel lines (figure 6) since they are spread over several pixels. Watershed regions belonging to these areas are merged together, seârating both theoretically adjacent roof facets. This could be avoided by assigning higher gradient valuesinthe middle of the roof tops.

Figure 7 (bi-plane building roofs) shows a consequence of an horizontal cut of the hierarchy without any geometrical criteria. If boundaries of watershed regions follow the roof top, the final delineation follows a curved line. An horizontal cut considers all regions at the same level without local properties. It is therefore planned to perform an automatic cut of the hierarchy following planarity criteria, that is performing an ascent from the leaf nodes (initial regions) to a node satisfying a planarity criterion (e.g. estimating a least square plane and defining a threshold on the sum of the residuals).

6 CONCLUSION

We have presented in this paper an approach for combining LIDAR 3D data and aerial images. At first, LIDAR data are processed to extract 3D planar primitives over buildings. This detector is based on a modified RANSAC paradigm guided with information concerning normal vectors of the scene. These primitives are then introduced in an image segmentation process based on a ascendant region merging scheme. From an initial segmentation (watershed on the image gradient), regions are merged under specific geomet-

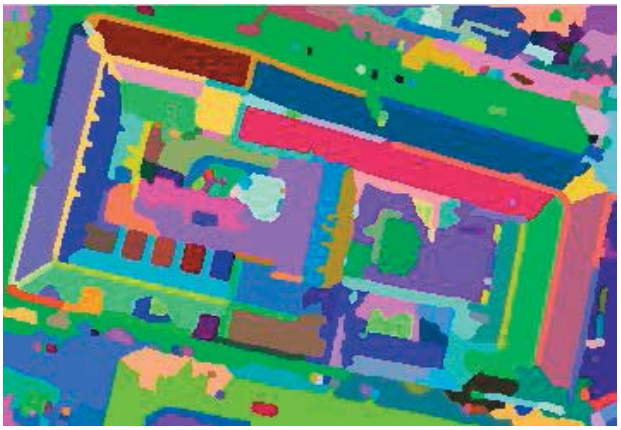


Figure 5: Final image partition of an ortho-image represented as a label image. The cut in the hierarchy was performed on visual criteria.

Figure 7: Details of the segmentation of figure 4

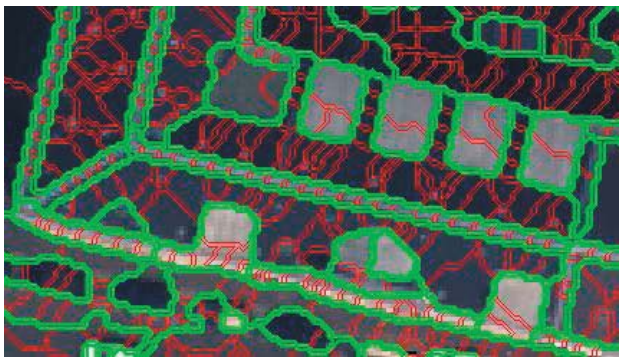


Figure 6: Details of the segmentation presented in figure 4. Gutters and roof tops are segmented independently of roof facets. We may notice that roof hyper-structures are sharply distinct from the building roof itself.

ric and radiometric constraints. The merging steps are performed on a region adjacency graph. A hierarchy is built whereon a cut provides a desired image partition. The future work consists of improving the general cost function as well as to perform automatic and non-horizontal cuts of the hierarchy with geometric considerations.

REFERENCES

Bretar, F., Chesnier, M., Pierrot-Desseilligny, M. and Roux, M., 2004. Terrain modeling ans airborne laser data classification using multiple pass filtering. Vol. XXXV part B, International Archives of Photogrammetry and Remote Sensing, Istanbul, pp. 314–319.

Bretar, F., Pierrot-Desseilligny, M. and Roux, M., 2003. Estimating intrinsic accuracy of airborne laser data with local 3D-offsets. Vol. XXXIV, International Archives of Photogrammetry and Remote Sensing, Dresden, Germany, pp. 20–26.

Filin, S., 2003. Analysis and implementation of a laser strip adjustment model. Vol. XXXIV-3/W3, International Archives of Photogrammetry and Remote Sensing, Dresden, pp. 65–70.

Fischler, M. A. and Bolles, R. C., 1981. Random sample consensus: A model fitting with applications to image analysis and automated cartography. Graphics and Image Processing 24(6), pp. 381–395.

Haris, K., Efstratiadis, S., Maglaveras, N. and Katsaggelos, A. K., 1998. Hybrid image segmentation using watersheds and fast region merging. IEEE Transactions on Image Processing 7(12), pp. 1684–1689.

Hartley, R. and Zisserman, A., 2000. Multiple View Geometry in computer vision. Cambridge University Press.

Hofmann, A. D., Maas, H. G. and Streilein, A., 2003. Derivation of roof types by cluster analysis in parameter spaces of airborne laserscanner point clouds. Vol. XXXIV, International Archives of Photogrammetry and Remote Sensing, pp. 112–117.

Horn, B., 1984. Extended gaussian images. PIEEE 72(12), pp. 1656–1678.

Maas, H. G. and Vosselman, G., 1999. Two algorithms for extracting building models from raw laser altimetry data. ISPRS Journal of Photogrammetry and Remote Sensing 54, pp. 153–163.

McIntosh, K. and Krupnik, A., 2002. Integration of laser-derived DSMs and matched image edges for generating an accurate surface model. ISPRS Journal of Photogrammetry and Remote Sensing (56), pp. 167–176.

McIvor, A. and Valkenburg, R. J., 1997. A comparison of local surface geometry estimation methods. Machine Vision and Applications 10, pp. 17–26.

Pottman, H., Leopildseder, S., Wallner, J. and Peternell, M., 2002. Recognition and reconstruction of special surfaces from point clouds. Vol. 34 part 3A, International Archives of Photogrammetry and Remote Sensing, Graz, pp. 271–276.

Rottensteiner, F. and Briese, C., 2003. Automatic generation of building models from lidar data and the integration of aerial images. Vol. XXXIV-3/W3, International Archives of Photogrammetry and Remote Sensing, Dresden, pp. 174–180.

Schenk, T. and Csatho, B., 2002. Fusion of lidar data and aerial imagery for a more complete surface description. Vol. XXXIV-3A/W3, International Archives of Photogrammetry and Remote Sensing, Graz, pp. 310–317.

Sithole, G. and Vosselman, G., 2003. Comparison of filtering algorithms. Vol. XXXIV-3/W3, International Archives of Photogrammetry and Remote Sensing, Dresden, pp. 71–78.

Sohn, G. and Dowman, I., 2003. Building extraction using LIDAR DEMs and IKONOS images. Vol. XXXIV-3/W3, Dresden, pp. 174–180.

Taylor, R., Savini, M. and Reeves, A., 1989. Fast segmentation of range imagery into planar regions. Computer Vision, Graphics and Image processing 45, pp. 42–60.

Vosselman, G. and Dijkman, S., 2001. 3D building reconstruction from point cloud and ground plans. Vol. XXXIV, International Archives of Photogrammetry and Remote Sensing, Annapolis, pp. 37–43.

Ward, J. H., 1963. Hierarchical grouping to optimize an objective function. Journal of the American Statistical Association 58(301), pp. 236–244.

Xu, G. and Zhang, Z., 1996. Epipolar Geometry in stereo, motion and object recognition. Kluwer Academic Publishers.