AUTOMATIC DETECTION OF BUILDINGS WITH RECTANGULAR FLAT ROOFS FROM MULTI-VIEW OBLIQUE IMAGERY

J. Xiao *, M. Gerke, G. Vosselman

Faculty of Geo-Information Science and Earth Observation, University of Twente, 7514AE Enschede, The Netherlands - (jing, gerke, vosselman)@itc.nl

Commission III,  WG III/4

KEY WORDS: Oblique Image, Multiple views, Building Detection, Plane Sweeping, Cross Correlation

ABSTRACT:
Automatic building detection plays an important role in many applications. Traditionally, buildings were detected from monocular or multi-view images, DEM or DSM, and laser scanner data. Oblique imagery is a relatively new data source with distinct advantages: it provides detailed information on facades and multiple views from various perspectives. In this paper we present exploration tests using only oblique imagery for the detection of buildings, mainly focus on rectangular flat roof type. It has two major stages: 1) generating robust facades from 3D lines extracted from multiple images; and 2) determining the height and roof outlines from a single facade of one building by plane sweeping and image segmentation. All tested buildings with rectangular flat roofs were successfully distinguished from other buildings, and their height and roof outlines were correctly detected. The novelty of this paper is the combination of geometric and radiometric approaches requiring only oblique imagery. This approach can be improved by adjusting the segmentation method and adding parameters for plane sweeping in order to be successful on buildings of other types.

1. INTRODUCTION

Automatic building detection is important in many applications, for instance map updating, city modelling and urban planning. Various data sources have been used for building detection, including single or multiple overlapped airborne images (Müller and Zaum, 2005; Karantzalos and Paragios, 2009), DSM or DEM (Ma, 2005; Lu et al., 2006), InSAR (Thiele et al., 2007) or laser scanning data (Oude Elberink and Vosselman, 2009), as well as combinations of these (Khoshelham et al., 2010). However, all these data sources provide only vertical scene information, thus making it difficult to distinguish single buildings when their roofs are connected and of homogenous appearance.

Nowadays, oblique images with large tilt angles are available. (Petrie and Walker, 2007). These provide abundant information on building facades. Objects are imaged with stereo overlap from multiple directions. This imagery also brings some challenges not present in the use of ortho-images. The first is the variable-scale geometry caused by the tilt angle (Grenzdörfner et al., 2008). Another is occlusion objects can be self-occluded on one or two sides, or be occluded by higher nearby objects.

The objective of this research is to develop a method to detect buildings from multi-view oblique images alone, without any other data sources. In this paper we only focus on buildings with flat roofs. The research questions are: 1) how to generate robust facades from line detection in oblique imagery; and 2) how to determine building height and roof outline in 3D. Some related work is firstly reviewed in section 2.

2. RELATED WORK

Multiple overlapped nadir images have long been popular for solving the problem of building detection (Roux and McKeown, 1994; Baillard et al., 1999; Kim and Nevatia, 2004), but off-nadir imagery has also been applied in a few studies. For example, buildings were detected using shadow and wall evidence in one oblique view (Lin and Nevatia, 1995), and large buildings were recognized from natural images (Malobabic et al., 2005). Terrestrial video is another source of image sequence for building detection (Tian et al., 2009).

Oblique imagery combines the advantages of multiple overlapped views and terrestrial imagery, but its applications are not so wide till now. The main application is texture extraction for 3D modelling (Frueh et al., 2004; Wang et al., 2008). A few studies have been carried out on making use of the inherent 3D information for dense matching (Le Besnerais et al., 2008; Gerke, 2009). Another study validating road data from imagery (Mishra et al., 2008). The accuracy of measurements from single-picture in a certain software using oblique imagery was assessed by Sukup et al. (2009).

3. METHODOLOGY

The methodology consists of two stages, each with sub-stages (Figure 1). In the first, the vertical facades of the buildings are generated from extracted 3D lines, which are considered robust but do not include all the lines in the scene. In the second stage, heights and outlines of the roofs are detected by plane sweeping and segmentation from the vertical facades.

* Corresponding author.
3.1 3D Line Extraction

The process begins with 3D line extraction from several oblique images taken from different perspectives. Line hypotheses are initially defined by stereo intersection of 2D straight lines extracted in the single images (Förstner, 1994). Since any intersection results in a 3D line, the main task is to exclude wrong matches. This is done by combining the stereo matches so that a particular 2D edge which is used for the 3D intersection with edges from other views must result in the same 3D line. The reliability of 3D line extraction is governed by requiring a minimum number of images to compute each line. Since the intersection results in infinite 3D lines, the last step consists in projecting the endpoints of the respective 2D straight lines onto the final 3D line and to use the extremas as 3D endpoints.

3.2 Facades Location

Extracted 3D lines are assumed to be reliable since they are validated by five or six images. According to their spatial distribution, lines are separated into groups and the facade is then fit within each group. Since the facades are vertical to the ground, their projections on the XOY plane in the 3D Cartesian coordinate system are used.

3.2.1 3D Line Grouping: In the real world, most vertical and horizontal planes above ground belong to buildings, thus the grouping of the lines is based on their coplanarity.

Vertical lines are separated from horizontal lines and partitioned into groups by minimum distance. For each group, the centroid on the XOY plane is determined by averaging the X and Y coordinates of all end points in the group.

Grouping of horizontal lines are carried out on their projections on the XOY plane, since we assume that all the walls are vertical. Lines are grouped only if their projections are close to each other and have small angle difference. The representative line of each group on the XOY plane is fit using least squares. A group may contain only one line.

3.2.2 Facade Fitting: Different strategies are applied to vertical and horizontal groups to generate vertical facades. The first strategy uses a vertical group and a nearby horizontal group to fit a facade. One single vertical group cannot define a facade independently since it can be from other objects such as the trunk of a tree or a pole. Additional information from horizontal groups near it would provide strong evidence for a facade. The location of the facade is defined by both the centre of the vertical line group and the representative line of the horizontal group. If they cannot fit perfectly with each other, a higher weight is allocated to the former one. The height of the facade is determined by the maximum height of all contributing lines. Two planes may be defined by the same vertical group as a vertical edge is the connection of two vertical facades.

The second strategy is used for the remaining horizontal line groups to define independent planes. Evidence provided by single-line and two-line groups is considered too weak to define a facade. Only groups containing three lines or more are selected. A new facade is built by extruding the representative line of a selected group on the XOY plane to the maximum height of line members.

3.3 Roof Location

Since the facade hypothesis is generated from at least three 3D lines, which is in turn validated by at least six images, it can be assumed to be rather robust. But the resulting disadvantage is that there is only one facade for each building in most cases. Therefore this facade must be used as the input for roof location.

Detection of builds from a single facade has two major problems: 1) to find out on which side of the facade the building is situated; and 2) to determine the actual height of the building. Plane sweeping in combination with cross-correlation are used to solve these problems.

Hypothesised heights for plane sweeping are defined by the searching step in the searching interval distributed evenly around the facade height. On each side of every hypothesised height, an average matching ratio \( R_{cc} \) is computed by segmentation of all images followed by calculation of the correlation coefficient for each image pair in the intersected segmented area. A curve representing \( R_{cc} \) values at all hypothesised heights for each side of the facade is drawn to determine the building roof type. A specific shape of the curve indicates the existing of one type of roof on that side, and the absence of building on the other side at the same time. The roof height is at maximum \( R_{cc} \) on the curve of the roof side.

3.3.1 Image Selection: Four images of the same building from different directions are automatically selected using their orientations among all available images. Although this slightly reduces amount of evidence, the four images form the strongest geometric relation. This also significantly reduces the calculation of \( R_{cc} \), whose computational time for \( n \) images is \( O(n^2) \).

3.3.2 Plane Sweeping: After image selection, hypothesised horizontal planes are tested at each hypothesised height. For each plane, two \( R_{cc} \) values are calculated for both sides by image segmentation and matching.

Image Segmentation: The facade line with one hypothesised height is projected onto each image. The Flood Fill function of OpenCV is used for segmentation, following the region-growing approach. On each side of the facade in each selected image, one seed point is defined near it. In order to avoid under-segmentation, the threshold for colour difference is set relatively low. To avoid over-segmentation, a range defined by
the length of the facade and a certain width perpendicular with the facade in the XOY plane is used. If the bounding box of the segmentation is within this range, new seed points will be appointed just outside the segmented area for further flood fill to enlarge the segmentation. All the segmentations are kept on one side of the facade.

Matching Ratio (Rcc) Calculation: This step is carried out at each side of the hypothesized facade. Because of the segmentation range, the result normally includes at least the whole roof. Therefore the calculation is performed on the intersection of the segments from four images, rectified according to the selected horizontal plane. An $n \times n$ window is used in the rectified images to calculate the correlation coefficient within each image pair. The matching ratio $Rcc$ is then defined as the average percentage of windows whose correlation coefficient is above a certain threshold $Tcc$.

Definition of Roof Side and Height: An $Rcc$ value is computed for each side of the facade for every hypothesized height. $Rcc$ vs. height is plotted for each side. If the target building has a rectangular flat roof and if the segmentation is on the correct side, an obvious peak at the correct height of the building can be expected in the curve. The ratio $Rr$ of the highest value of one curve to the median value is used to identify such buildings. In the case that there is no such trend on either side, the building is assumed to have another roof type.

3.4 Roof Outlining

After known one roof edge from the facade, the roof side and height, we try to extract robust roof outline so as to make the cube of the building.

3.4.1 Line Extraction: In each of the selected image, line segments are extracted within the region by image segmentation at the correct side and height. The region is extended outwards a bit to make sure that the whole roof is covered.

3.4.2 Line Validation: The roof outline is determined in 3D object space. Extracted lines from the four images are projected into the object space at the correct height. Each 3D line from one image is verified by lines from other images. Considering the possible incomplete line extraction owing to occlusion or poor radiometric contrast, a line is validated if verified from at least two images.

3.4.3 Outline Formation: In most cases, the valid roof lines are in a sparse organization – they are not connected to each other or there are small lines from the tiny structures on the roof, so the last step is to formalize the outline. Since the paper merely aims at rectangular flat roof, a rectangle is used to fit the outline by adjusting its direction, length and width.

4. EXPERIMENTS

4.1 Study Area and Data

The study area is located in the southeast of Enschede, The Netherlands. It can be divided into two regions: one with isolated buildings while the other with regular residential houses connected together. The two region included buildings with rectangular flat roofs which may also have various tiny structures, flat roofs of other shapes, and gable roofs which may be multi-colored.

The oblique imagery used in this research (Table 2) was collected in February 2007 by Pictometry through BLOM Aerofilms (Pictometry, 2010). Four oblique images looking forward, backward, left and right were taken at one time by four small frame cameras. For one object, four to eight images from different perspectives were available. All the images used in the experiment were oriented using the method described by Gerke and Nyaruhuma (2009). The RMSE at check points in object space was around 20cm for all three components after the self-calibration bundle adjustment.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flying height (m)</td>
<td>920</td>
</tr>
<tr>
<td>Baseline (m)</td>
<td>400</td>
</tr>
<tr>
<td>Tilt angle (degree)</td>
<td>50</td>
</tr>
<tr>
<td>Directions</td>
<td>Forward, backward, left, right</td>
</tr>
<tr>
<td>Focal length (mm)</td>
<td>85</td>
</tr>
<tr>
<td>Pixel size ($\mu$m)</td>
<td>9</td>
</tr>
<tr>
<td>GSD fore-/background (cm)</td>
<td>10 - 16</td>
</tr>
<tr>
<td>Sensor size ($mm \times mm$)</td>
<td>$36 \times 24$</td>
</tr>
</tbody>
</table>

Table 2. Parameters of images from Pictometry

LiDAR data used for the verification of building height was collected at 13/14 March 2007, almost the same time with oblique images. Its average point density is 20 pts/m² (Vosselman, 2008), with a height accuracy of 10 cm by point measurement on solid objects.

4.2 Test on Buildings

Samples of in total 17 buildings with all roof types were collected, including nine buildings with rectangular flat roofs, two with non-rectangular flat roofs and six with gable roofs.

Based on pre-tests, an $11 \times 11$ window was chosen for calculating the correlation coefficient and $Tcc$ was set to 0.6 to compute the matching ratio. To significantly reduce the number of calculations, the coarse searching step length for plane sweeping was set to 0.5 meters in the search interval of 10 meters around the initial facade height. This step aimed at defining the roof type, roof side if it was a rectangular flat roof, as well as the approximate height for the later refinement with a 0.05 meter step length. A threshold $Tr$ for $Rr$ was trained by the test data to identify the rectangular flat roofs.

Typical buildings with rectangular flat roofs in both regions were chosen for the experiment. Some were tall and large independent buildings, while some of them were low and small residential buildings. Their heights measured from LiDAR data were taken as the true values for the later assessment for the heights decided by plane sweeping.

The dominant roof type besides rectangular flat roof was gable roof, therefore some of these were chosen to test the ability of the methodology to distinguish rectangular flat roof from gable roof buildings. Another two buildings with pentagonal flat roofs were also chosen for the test. Since the segmentation approach was designed particularly for rectangles, its behaviour on other shapes could be tested on these two pentagonal flat roof buildings.
5. RESULTS AND DISCUSSION

5.1 Result from 3D Line Extraction and Facade Location

3D lines were extracted from eight images looking at the same region from different directions and validated by at least six images. An example region is shown in Figure 3. In this region, most of the extracted lines are horizontal lines and a few are vertical, but none are in other directions. This result is consistent with the real situation of the region since buildings were the only objectives and all the building edges were either horizontal or vertical.

Extracted lines distributed mostly near the top of buildings. This is because lines at the lower part were occluded from some directions, thus it was not possible to obtain six validation images. The dense horizontal lines on some facades were probably induced by the complex horizontal structures.

Using these extracted 3D lines, the output of the facade location step was the vertical facade planes (Figure 3(c)). The length of each facade was determined by the end points of its grouped line members, and its height was defined by the maximum height of the members. Almost all the dominant facades were correctly located except three, which are marked in the dashed circles. Those wrong facades have lower heights than the others. The 3D lines leading to them probably came from the edge between the vegetation and the pavement at the base of the buildings. They can be culled either by setting a certain threshold of the heights or by the later process of building identification.

In the example area, only one facade was allocated for each of the eight buildings (ten in total). The detected facades for the four high buildings near the road and the two rectangular flat roof buildings behind were facing south. This could be because of the complex structures on their south facades. More lines were detected from the complex structure, thus more 3D lines on the facades were extracted. The failure of line detection on facades facing other directions may be due to the low contrast in some images.

5.2 Result from Roof Location

A histogram presenting $R_{cc}$ was plotted after coarse plane sweeping step for each of the tested buildings. By the histogram, rectangular flat roof buildings can be clearly identified from others. Then their heights were refined by plane sweeping with 5cm step length, and later verified by the heights measured from LiDAR data.

5.2.1 Preliminary Type and Height Definition: Examples of a rectangular flat roof building and a gable roof building are shown in Figure 4. For each of them, a histogram of the matching ratio $R_{cc}$ on both sides was plotted. As shown in Figure 4(c), there was a high peak on side one, identifying a rectangular flat roof, but no obvious peak on side two due to the lack of building. The histogram from the gable roof building (Figure 4(d)) shows an irregular pattern with no clear peak on either side. Therefore, the histogram was able to identify the type and location of the rectangular flat roof building.

General result was shown in Figure 5 for all the 17 tested buildings. Not all the flat roof buildings among them had the homogeneous roof plane as the one in Figure 3(a). Some of them had different colour patches or textures on top, one with many small dorms on top and the ninth building in the test having four small towers on the roof. Moreover, the tested buildings also contained some ones having big ratios of length to width.

As shown in Figure 5, $R_r$ values of all flat roof buildings with rectangle shape were much higher than others, so that they can be successfully identified. According to the test data, $T_r$ can be set from 3.5 to 5, but we chose 5 due to the preference of a robust detection on target building type. The wrongly excluded rectangular flat roof buildings should be able to identify in the further detection on the left buildings.
Refinement of Roof Height: For each identified rectangular flat roof buildings, its height was refined by plane sweeping with shorter searching step, and then compared with its height measured from LiDAR data.

<table>
<thead>
<tr>
<th>ID</th>
<th>Coarse Height (m)</th>
<th>Fine Height (m)</th>
<th>LiDAR Height (m)</th>
<th>Difference (Coarse – LiDAR) (m)</th>
<th>Difference (Fine – LiDAR) (m)</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>61.06</td>
<td>61.11</td>
<td>60.66</td>
<td>0.40</td>
<td>-0.45</td>
<td>0.19</td>
</tr>
<tr>
<td>2</td>
<td>57.28</td>
<td>57.23</td>
<td>57.20</td>
<td>0.08</td>
<td>0.03</td>
<td>0.03</td>
</tr>
<tr>
<td>3</td>
<td>65.07</td>
<td>64.97</td>
<td>64.97</td>
<td>0.10</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>4</td>
<td>53.40</td>
<td>53.55</td>
<td>53.60</td>
<td>-0.20</td>
<td>-0.05</td>
<td>-0.15</td>
</tr>
<tr>
<td>5</td>
<td>54.72</td>
<td>54.67</td>
<td>54.60</td>
<td>0.12</td>
<td>0.07</td>
<td>0.15</td>
</tr>
<tr>
<td>6</td>
<td>50.14</td>
<td>49.99</td>
<td>50.20</td>
<td>-0.06</td>
<td>-0.21</td>
<td>-0.15</td>
</tr>
<tr>
<td>7</td>
<td>51.44</td>
<td>51.34</td>
<td>51.52</td>
<td>-0.08</td>
<td>-0.18</td>
<td>-0.15</td>
</tr>
<tr>
<td>8</td>
<td>76.00</td>
<td>76.20</td>
<td>76.25</td>
<td>-0.25</td>
<td>-0.05</td>
<td>-0.15</td>
</tr>
<tr>
<td>9</td>
<td>59.38</td>
<td>59.38</td>
<td>59.53</td>
<td>-0.15</td>
<td>-0.15</td>
<td>-0.15</td>
</tr>
</tbody>
</table>

Table 6. Comparison between coarse heights, refined heights and heights measured from LiDAR data.

As predicted, pentagonal flat roofs could not be distinguished from gable roofs. This was probably due to the background included in the rectangle segments outside the pentagon. However, in the histogram of a pentagonal flat roof (not shown in this paper) there was a rise in the curve, but wider and lower than the peak for a rectangular flat roof. Therefore, detection of buildings of this type could be improved by exploring the use of the histograms, or by adjusting the segmentation to different shapes.

6. CONCLUSIONS AND FUTURE WORK

The method shown in this paper requires no previous knowledge to detect buildings. By combining geometric and radiometric features of multi-view oblique imagery, outlines of simple cubic buildings can be successfully detected. The height accuracy of the detected rectangular flat roof buildings is 0.2m, which is acceptable.

Detection of buildings with flat but not rectangle roofs can be achieved by looking for new clues on the histogram from plane sweeping, or by adjusting the shape of segmentation. Referring to the work by Baillard et al. (1999), buildings with gable roofs or inclined roofs will be tested with improved plane sweeping strategy. Instead of using only height as the parameter, angle will also be used.

ACKNOWLEDGEMENT

I would like to thank BLOM Aerofilms for providing the Pictometry dataset. Then I would like to thank Adam Patrick Nyaruhuma for providing the code for image matching and David Rossiter for the revision. I would also like to thank the anonymous reviewer for their precious comments.

REFERENCES


