Abstract

In this paper, we propose a complete semi-automatic processing chain able to provide, from a couple of high-resolution optical and SAR images, a simple three-dimensional reconstruction of the buildings in urban scenes. The chain is decomposed into four steps: building boundary extraction, footprint projection and registration, height estimation and building validation. Each of these steps is described from a methodological point of view and applied on a scene of interest on Quickbird and TerraSAR-X images. Good results of building detection and height retrieval are obtained and a satisfying final 3D reconstruction of buildings is provided.

1 Introduction

Since a few years, new approaches, exploring the high detail level characterizing high-resolution optical and SAR data provided by current spaceborne sensors, have been proposed for object extraction and reconstruction in urban areas. Especially, challenges are born in the field of building detection and building height retrieval for 3D reconstruction of urban scenes.

Some techniques such as stereoscopy, radargrammetry and interferometry or simulation-based methods [1], have been applied in this framework and have given satisfying results of building extraction and height estimation. Nevertheless, some difficulties are still present: optical couple availability, need of auxiliary input data, incomplete or noisy results from SAR images...

In this paper, a new method is developed in the frame of optical - SAR data fusion [2]. A complete and operational semi-automatic processing chain, able to provide a simple 3D reconstruction of buildings in large urban scenes, is presented. The process takes only as input one HR optical image and one HR SAR image of the same area and uses complementary information extracted from both data. A simple parallelepiped is used as 3D building model for reconstruction (rectangular footprint, flat roof).

The chain can be decomposed into four main steps, described in Sections 2, 3 and 4. In Sections 2 and 3, the methodology previously proposed in [2] for building boundary extraction from monoscopic optical image and for footprint projection and registration into SAR data, is reminded and illustrated on a Quickbird and TerraSAR-X couple of an urban scene. In Section 4, the main contribution of this paper is largely presented. A new approach for potential building validation and height retrieval from SAR data, based on the combination of two complementary criteria, is developed. First, some theoretic and practical considerations about the strategy are proposed. Then, the process is applied on the whole studied scene and the results obtained for the reconstruction of buildings are analyzed.
differently extracted rectangular subparts of a same building, can be performed by taking into account juxtaposition aspects and similarities of orientation and dimensions. The whole extraction process has been applied to a Quickbird urban scene (Marseille - France), that will be used throughout this paper to illustrate our results. After this first step of the global processing chain, a fine binary map precisely referring to the bidimensional boundaries of the extracted footprints is obtained (see Figure 1).

After this first step of the global processing chain, a fine binary map precisely referring to the bidimensional boundaries of the extracted footprints is obtained (see Figure 1). The post-processing has then been applied to footprints E and F, fused into a larger one (footprint D).

Figure 1: Result of the building extraction process on an optical urban subscene (Quickbird sensor ©DigitalGlobe, resolution: 0.6 m, Marseille - France).

3 Footprint Projection and Registration into SAR Data

To combine information issued from both optical and SAR images, we have to perform a step of projection and registration between homologous building features. In our case, the extracted rectangular optical footprints are projected into the SAR image and matched with their corresponding features, i.e. with the bright linear SAR echoes issued from double ground / wall reflections. The original proposed method of projection and registration, presented in [2], has been applied on the studied scene. A first phase of projection is realized by using the available physical parametric joint model of projection, and a second phase of registration refinement is performed through a criterion optimization on the SAR image. After this second step of the global processing chain, a fine and satisfying adjustment (at building scale) of the projected footprint location in the SAR image is obtained.

4 Validation and Height Retrieval by Fusion with SAR Features

4.1 Objective and Global Strategy

The objective of this Section is to explore complementary information provided by SAR data, to jointly validate building detection and estimate their height. A new automatic approach, using the combination of a statistical region-based criterion and a radiometric contour-based criterion, depending on the building height and computed on building SAR areas, is proposed. The statistical global criterion refers to geometric constraints relative to positions and dimensions of homogeneous characteristic building SAR areas, while the radiometric local criterion takes into account an a priori condition that can be introduced by measuring a local contrast on discontinuity areas. The methodology is based on a process in three phases: First, the statistical criterion is optimized to provide a set of candidate building heights. Then, the radiometric criterion is used to retrieve the final estimated height, by selecting the one corresponding to the best value of this second criterion. Finally, a last step of thresholding on this best radiometric contrast, allows to validate or not the building.

4.2 Phase 1: Statistical Criterion for Candidate Building Height Generation

4.2.1 Optimization Scheme

The main idea is to delimit on the SAR image the position of building characteristic SAR regions, inferred by the optical footprint and an associated height. We propose to use, on the SAR image, a statistical criterion, based on the partitioning of the SAR image in different regions (layover, single roof, shadow...) depending on the supposed height. The criterion optimization allows us to retrieve a set of most likely building heights. The different areas composing the partitioning are defined according to a classical modeling of the geometric building signature on SAR images and can be listed as follows: the bright "double echo area" corresponding to a double ground / wall reflection, the "layover area" corresponding to the simultaneous backscattering from ground, roof and frontages, the "single roof area" where only the roof is responding, the "shadow area" where no object is responding and the "background area". As described above, the method is based on the global scheme of: "Height Hypothesis - Test Partitioning Generation on SAR Data - Statistical Criterion Optimization". For each potential building footprint, detected on the optical image, projected and registered into SAR data, the following steps are applied.

1. First, the position of the four building corners are deduced in the SAR image from the rectangular footprints. The building subscene, on which the criterion will be computed, is defined by extending the footprint bounding box.

2. Then, the scheme for the generation of candidate building heights is performed:
   - A height test value belonging to an interval of potential building heights is iteratively chosen.
   - The partitioning of the subscene into the different characteristic areas is computed in function...
of the test height, the footprint location and dimensions and in function of some known SAR acquisition system parameters (incidence angle and range spacing), by using a geometric SAR model of projection.

- The statistical criterion is then computed on the test partitioning, as described in the following, by using SAR intensity of pixels belonging to the different areas.

3. Finally, the generation of candidate building heights is done by analyzing the curve representing the evolution of the criterion, in function of the height test values. When a representative local optimum clearly appears, a strong geometric adequation exists between the predicted signature and the real one: the associated height test value is thus conserved as candidate building height for the second phase of the process. Let us remark that it can be a local or global optimum, depending on the scene complexity.

4.2.2 Definition of the Statistical Criterion

In order to test the adequation between a test partitioning and the reality imaged on SAR data, we propose to use as statistical criterion the Log-Likelihood, computed on the individually extracted building SAR subscenes.

The global Log-Likelihood is defined as the sum of the elementary Log-Likelihoods, computed on the different areas in the subcene. For each SAR area, the intensity distribution is classically modeled by a Gamma density function defined by two parameters. The analytical expression of the Log-Likelihood of area \( i \) is thus given as follows:

\[
LL_{area}(i) = \frac{N_i(L_i \log(L_i) - \log G(L_i))}{\mu_i} - \sum_{j \in i} x_j - N_i \log(\mu_i) + (L_i - 1)(\sum_{j \in i} \log(x_j) - N_i \log(\mu_i))
\]

where \( N_i \) is the number of pixels in the area \( i \), \( x_j \) is the intensity of pixel \( j \) and \( \log G \) is the Log-Gamma function.

Both \( (\mu_i, L_i) \) parameters are locally estimated on each area by the Maximum Likelihood method.

Our final statistical criterion is defined as the opposite of the global Log-Likelihood and corresponds to an energetic cost, that has to be minimized.

The candidate heights issued from the optimization correspond to the most likely heights and are defined as:

\[
h_{cand} = \arg\min_{local} \left( - \sum_{i \in \text{areas}} LL_{area}(i) \right)
\]

4.3 Phase 2: Radiometric Criterion for Final Building Height Selection

To improve this statistical criterion, we propose to combine it with a radiometric criterion, testing the presence of predicted discontinuities. In particular, the presence of an edge between the "background area" and the "layover area", whose location depends on the height test value, is tested by computing the local contrast between both areas. This criterion is defined as the mean amplitude ratio of pixels belonging to narrow bands on both sides of the discontinuity. Given the a priori knowledge that the "background area" is darker than the "layover area", this ratio has to be smaller than one, and, when the partitioning of the subscene generated from a test height matches perfectly with the SAR image, this ratio has to be minimal.

We propose to proceed as follows:

1. First, this ratio, denoted by \( r(h) \), is calculated for each height test value \( h \), belonging to the interval of potential building heights.

2. Then, for each candidate building height \( h_{cand} \), the minimal ratio value, obtained for a test height belonging to a close neighbourhood around \( h_{cand} \), is stored. This local minimum ratio value, denoted by \( r(r_{cand}) \), can be obtained for a slightly different height than \( h_{cand} \), that is denoted by \( r_{cand} \) in the following.

3. Then, the candidate height \( h_{cand} \) running to the best minimal contrast \( r(r_{cand}) \) is selected.

4. At the end, the final estimated height \( h_{est} \) is defined as the mean between the selected height \( h_{cand} \) and the associated height \( r_{cand} \), issued from both criteria.

4.4 Phase 3: Building Validation

We finally perform a thresholding on the best radiometric contrast \( r(r_{cand}) \), in order to validate or not the building presence. If the test is not passed, the building is rejected.

After this third step of the global processing chain, a grey level map indicating the relative estimated heights of validated buildings is provided.

Finally, as our algorithms are based on a simple 3D parallelepipedic building model, a 3D view of the buildings composing the scene, can be easily generated by combining the planimetric information issued from the boundary map and the altimetric information issued from the height map of validated buildings.

4.5 Application on a Real TerraSAR-X Urban Scene

The proposed approach has been tested on the studied urban scene (TerraSAR-X image, Marseille - France).
The incidence angle value is 31°. The interval of heights is 3 m to 15 m. The close neighbourhood around candidate heights is described by a 5 meter large window. The thresholding coefficient for the local contrast value $r(h_{\text{ratio}})$ is fixed to 0.8 in order to prevent this thresholding from being too much strict (indeed, a light thresholding can appear preferable as the contrast value presents a high variability due to noise and punctual disturbing targets on real data).

Table 1 summarizes the results obtained for each extracted building footprint, composing the studied scene. The first column refers to the building footprint label. The second one indicates the value of the selected candidate height $h_{\text{cand}}$. The third one indicates the value of the height $h_{\text{ratio}}$ providing an optimal value $r(h_{\text{ratio}})$ (given in the fourth column) of the radiometric criterion. The two last ones indicate the final estimated height value $h_{\text{est}}$ and the decision resulting from the building validation step.

Figure 2 (a) represents the obtained grey level map of building heights. It indicates the relative retrieved heights of buildings, whose validation has been performed by taking into account the previously exposed combined approach. Figure 2 (b) shows the final 3D view of buildings composing the studied scene, that has been visually enhanced by integrating the SAR image texture.

<table>
<thead>
<tr>
<th>Label</th>
<th>$h_{\text{cand}}$</th>
<th>$h_{\text{ratio}}$</th>
<th>$r(h_{\text{ratio}})$</th>
<th>$h_{\text{est}}$</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>5 m</td>
<td>7 m</td>
<td>0.60</td>
<td>6 m</td>
<td>Validated</td>
</tr>
<tr>
<td>B</td>
<td>9 m</td>
<td>7 m</td>
<td>0.72</td>
<td>8 m</td>
<td>Validated</td>
</tr>
<tr>
<td>C</td>
<td>5 m</td>
<td>5 m</td>
<td>0.64</td>
<td>5 m</td>
<td>Validated</td>
</tr>
<tr>
<td>D</td>
<td>7 m</td>
<td>9 m</td>
<td>0.75</td>
<td>8 m</td>
<td>Validated</td>
</tr>
<tr>
<td>G</td>
<td>9 m</td>
<td>9 m</td>
<td>0.18</td>
<td>9 m</td>
<td>Validated</td>
</tr>
<tr>
<td>H</td>
<td>4 m</td>
<td>4 m</td>
<td>0.61</td>
<td>4 m</td>
<td>Validated</td>
</tr>
<tr>
<td>I</td>
<td>6 m</td>
<td>6 m</td>
<td>0.72</td>
<td>6 m</td>
<td>Validated</td>
</tr>
<tr>
<td>J</td>
<td>8 m</td>
<td>7 m</td>
<td>0.73</td>
<td>8 m</td>
<td>Validated</td>
</tr>
<tr>
<td>K</td>
<td>12 m</td>
<td>14 m</td>
<td>1.2</td>
<td>-</td>
<td>Rejected</td>
</tr>
</tbody>
</table>

Table 1: Validation status and estimated heights.

![Height map](image1) ![Final 3D reconstruction](image2)

Figure 2: (a) Result of building validation and height estimation process. (b) Result of the 3D reconstruction obtained for the studied urban scene (TerraSAR-X sensor ©Infoterra, resolution: 1.1 m, Marseille - France).

4.6 Analysis of the Results

Among the 9 status delivered at the end of the process, 7 of them are true. The building K has been correctly designed as false alarm and corresponds indeed to an homogeneous rectangular parcel on the ground. Among the 8 validated buildings, 6 of them correspond to a correct building detection, whereas two of them (buildings C and H) are false detections. The height retrieval process has provided relevant estimations of the building heights, that are in accordance with a visual estimation of the relative true building heights, based on ground truth photographies. A satisfying 3D reconstruction of the scene has thus been achieved. With such a height estimation process, we have taken advantage of the robustness characterizing the global statistical criterion and of the precision and discrimination ability characterizing the local radiometric criterion.

4.7 Potential Improvement

An improvement aiming to reduce the rate of false alarms could be achieved by computing a set of confidence scores, defining the reliability of the final estimated heights and characterizing the behaviour of both criteria around these heights. For instance, these scores would be relative to the criterion value or to the peak shape, weight and precision, visible on criteria evolution curves. A thresholding on the combined global confidence score could thus help us to improve the validation step by deleting false alarms, designed by a weak confidence score.

Such an improvement, applied to the studied scene, would help us to remove the two false alarms inducing by the potential footprints C and H.

5 Conclusion and Future Work

A complete semi-automatic processing chain, providing relevant 3D information of buildings from a couple of HR optical and SAR images, has been proposed. It has been shown that the process is able to efficiently deal with different building size. A satisfying building detection has been obtained and an approach combining statistical and radiometric criteria has permit to achieve good results for building height estimation on the scene. In further works, the complete methodology could be extended to more sophisticated buildings, in order to deal with the reconstruction of complex urban scenes.

References
