This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier’s archiving and manuscript policies are encouraged to visit:

http://www.elsevier.com/copyright
Populating a building Multi Representation Data Base with photogrammetric tools: Recent progress

Benoit Frédéricque\textsuperscript{a,*}, Sylvie Daniel\textsuperscript{a}, Yvan Bédard\textsuperscript{a}, Nicolas Paparoditis\textsuperscript{b}

\textsuperscript{a} Université Laval, SCG, Canada
\textsuperscript{b} IGN, Laboratoire MATIS, France

Received 29 June 2007; received in revised form 18 January 2008; accepted 21 January 2008
Available online 23 May 2008

Abstract

More and more frequently, the populating of 3D MRDB’s (Multi Representations Data Base) is required to support advanced cartographic applications and advanced geospatial decision analysis. We previously proposed a manual photogrammetric process based on the Multi Representation Acquisition Pattern (MRAP) concept to tackle simultaneously Fine Level Geometries (FLG) extraction and Coarse Level Geometries (CLG) extraction. This paper presents our progress, from an automation standpoint, regarding our approach to populate MRDBs containing building geometries through a photogrammetric multiple representations acquisition process. Two new algorithms dedicated to Multi Representation (MR) acquisition are introduced and constitute two contributions. These combine passive imagery and Digital Surface Model analysis in order to address automation issues. The first algorithm allows for the automatic determination of the MRAP parameters from a single click initialization. The second algorithm aims at supporting part of the MR acquisition when MRAP are not relevant and allows the automatic building footprints extraction. This paper describes the project motivation and its actual progress. It is divided into four parts. The first part concerns the MR data acquisition specifications. A description of the manual prototype is provided and the MRAP concept is described. In the second part, the first algorithm allowing the automatic determination of the MRAP parameters is introduced. The third part is dedicated to the review of the automation performances through the study of three test sites. Finally, our second algorithm, allowing the automatic building footprints extraction is introduced with preliminary results.

Keywords: Building extraction; Multi representation Data Base; Multi representation acquisition pattern; Graph matching; A priori knowledge

1. Introduction

Emerging web-mapping applications and SOLAP (Spatial On-Line Analytical Processing) applications (Bédard et al., 2007) have become increasingly demanding from a map production standpoint. On-demand map production with highly customizable capabilities is needed. In fact, these web-mapping and SOLAP applications need to manage the map contents at the instance level (Bernier et al., 2005). Instead of having the abstraction levels provided in the maps stored by dataset layers, each geographical object should manage its own abstraction levels. The map user interface should also support navigation...
operations through the different abstraction levels, like drilling-up, drilling-down or drilling across. Producing this kind of drillable map is a complex task. This has motivated a lot of research work, particularly in the field of cartographic generalization (Cecconi et al., 2002; Mackaness et al., 2007). In fact, using a unique detailed data source (preliminarily extracted) to generate on-the-fly (using generalization) simplified geometric representations of geographic objects would fulfill these needs. Despite a lot of research work, generalization is still time consuming and involves a large amount of human intervention. An alternative to on-the-fly generalization is using Multi Representations Data Bases (MRDB), that store an explicit link between various geometries of geographical objects. Storing and linking several geometric representations of each geographical object in a geospatial database allows fulfilling the needs described previously. It is worth noticing that CityGML \(^1\) defines and manages the most relevant topographic objects in cities at several abstraction levels. In this information model, the links between the geometries corresponding to the different abstraction levels, the Levels Of Details (LODs), are also explicit. The use of MRDB to generate map-on-demand has motivated a lot of research (Weibel and Dutton, 1999; Cecconi et al., 2002; Bédard, 2004; Mackaness et al., 2007; Bernier and Bédard, 2007).

When populating a MRDB, three approaches can be used to extract and link fine and coarse geometries: (1) Fine Level Geometry (FLG) extraction and generalization (the Coarse Level Geometries (CLG) are deduced from FLG using the generalization process) (Cecconi et al., 2002; Mackaness et al., 2007), (2) geometric and semantic matching of different sources at different scales (FLG and CLG are already available, the link is built through a matching process) (Bernier and Bédard, 2007; Otteau et al., 2006), (3) multi representations (MR) data acquisition (simultaneous acquisition and link of CLG and FLG). Obtaining multiple geometric representations to populate MRDBs implies several difficulties from the automation standpoint and is therefore a cost issue. Despite important research work to automate them, a large amount of human interventions are still needed for the first two approaches. More detailed information on the above topics and reviews of related work can be found in Frédéricque et al. (2005) and Frédéricque et al. (2008). The third approach, which consists in defining several abstraction levels during the acquisition step (Fine and Coarse), is very recent. We are, to our knowledge, the only research group to work on such an approach. Although FLG and CLG extractions are both extremely complex tasks that, today, cannot be entirely automated, we think these two major stages of populating MRDBs present some similarities. Tackling them simultaneously using a photogrammetric data acquisition approach could be advantageous. In Frédéricque et al. (2005), Frédéricque et al. (2008), we proposed a multi representation data acquisition framework and a manual photogrammetric tool to perform MR acquisition. The proposed framework was designed based on our review of existing works addressing (1) MRDB population and (2) buildings extraction through photogrammetric process. We proposed a semi-automatic strategy involving human interventions at the beginning of the acquisition process. The purpose of the human interventions is to introduce a priori knowledge useful for both FLG and CLG extraction. The MRAP concept, formalizing a priori knowledge about building geometries, was introduced to support MR acquisition. This links two existing concepts, parametric models and geometric patterns, introduced respectively in the photogrammetric and generalization communities. The MRAP concept was implemented in our manual photogrammetric tool. Since MRAP relies on parametric models, this approach is not relevant in all configurations. In other words, all the buildings cannot be described a priori with MRAP. However, we think that they are relevant in most cases, particularly in a North American context, and that they can consequently play a major role in a MR acquisition system. More flexible approaches have to be considered to complete a MR acquisition system.

This paper describes our recent progress from an automation standpoint. We are addressing contexts where MRAP is relevant and not. Two new algorithms, using aerial imagery and Digital Surface Models (computed from aerial imagery), are introduced. They represent two innovative contributions to the automatic building extraction field of work. The first can be applied when the MRAP concept is relevant and allows for the automatic determination of the MRAP parameters from a single click initialization. The second algorithm aims at supporting part of the MR acquisition when MRAP are not relevant to tackle FLG and CLG extraction. This second algorithm allows for automatic building footprint extraction and representation as a combination of rectangles.

\(^1\) CityGML is an OGC adopted best practice paper providing a common semantic information model for the representation of 3D urban objects, \url{http://www.citygml.org/}. 
This paper is divided in four sections. The first section concerns the MR data acquisition specifications and our precedent works. This section is proposed in order to better understand the approach so far, as some results have only been published in French (i.e. the manual MR acquisition prototype). In this section, the manual prototype and the MRAP concept are described. In the second part, the first algorithm allowing enhancement of the level of automation of the MR acquisition using MRAP is proposed. A brief review of related works is also provided. The third part is dedicated to the assessment of the automation performance using MRAP and three sites have been used to test the algorithm. Finally, our second algorithm as well as preliminary results are introduced. This second algorithm aims at supporting part of the MR acquisition when MRAP are not relevant through the automatic building footprint extraction.

2. Multi representation data acquisition: existing tools

2.1. Goal

Given a territory and a range of abstraction levels, the photogrammetric MR data acquisition process aims at extracting and linking geometric representations of geographic objects. These geometric representations can be described by traditional CAD structures, such as line, polyline, point, etc. or by more expressive topological GIS structures, such as surfaces, multi-surfaces, and so on (a more detailed description of this structure can be found in the ISO-TC19125-1 standard). As suggested by our research group (Sabo, 2007), it can also be described using geometric patterns whenever possible.

The geometric pattern concept as well as its use to populate MRDB with a photogrammetric approach using MRAP is introduced below. More detailed information on these concepts can be found in Sabo et al. (2005a,b), Cardenas (2004) and Frédéricque et al. (2008).

2.2. Geometric pattern concept

2.2.1. Definition

A geometric pattern is defined as “a geometric object with basic geometric characteristics that are typical and representative of a large number of occurrences of a mapping feature-type or of a geometric primitive and that is able to adapt itself to the geometry of these occurrences of object at different scales and that can be reused several times” (Bédard, 2004; Cardenas, 2004). A geometric pattern consists of primitives that correspond to the indivisible atomic elements it is made of.

2.2.2. Purpose

Geometric patterns can be used to describe the geometries of an object stored in a database (DB). The use of geometric patterns enhances the integrity of a DB by reducing information redundancy. Moreover, the use of geometric patterns provides several advantages from the generalization standpoint. In Sabo (2007), the author suggests the combination of geometric patterns with simple generalization algorithms and generalization constraints to obtain Self-Generalizing Objects (SGO). SGO can be abstracted as software agents that perform generalization operations through the manipulation of the geometric pattern implantation parameters. For example, an exaggeration operation, applied on a rectangular building that is defined with a geometric pattern, is simply performed by increasing the building width and length parameters. A SGO can produce geometric representations at several arbitrary abstraction levels (which is impossible when using only MR stored in DB). It can also take into account the relations between different geographic objects. The reader should refer to Sabo (2007) for more details about SGO, their use and the quality of the resulting cartographic representations.

2.2.3. Links with the parametric model concept

The literature review we conducted reveals proximity between the 2D\(^2\) geometric pattern concept used to define CLG and the parametric models concept defined to capture FLG. Geometric patterns have been introduced by cartographers and database specialists for generalization purposes while the parametric model has been introduced to support photogrammetric data capture. Despite these two different origins, these concepts display strong similarities. In fact, both concepts use a library of a priori defined shapes to describe the geometries of geographical objects. The geometry description is carried out by defining specific implementation settings for these shapes. Like geometric patterns, parametric models can be used to reduce redundancy in a geospatial database and to ensure MRDB integrity. Several differences still remain between these two concepts. The differences can be

---

\(^{2}\)We consider a geometric object as 3D as soon as its geometry defines a volume and as 2D when its geometry is only defined in a plane.
Fig. 1. Basic volumetric primitive types: (a) flat roof, (b) pent roof, (c) gable roof, (d) hip roof.

categorized according to the three following aspects: (1) their dimensions (2D for the geometric patterns and 3D for the parametric models), (2) the addressed levels of abstraction (geometric patterns address CLG definition while parametric models address FLG), (3) the priority given to the geometric precision (the geometric patterns defined by Sabo (2007) voluntarily favor the generalization speed and data volume reduction while the parametric models favor the geometric fidelity).

2.3. MRAP concept

As mentioned before, our goal consists in populating MRDB using, when possible, geometric patterns and parametric models. Following such an approach, populating MRDBs requires, for each level of abstraction, to identify, implement, and link geometric patterns and parametric models corresponding to geographical objects. Eventually, generalization algorithms and generalization constraints can be defined to manage relations between spatial objects. This can be long and monotonous if done manually. To facilitate this process, we introduced the concepts of Multi Representation Acquisition Pattern (MRAP).

<table>
<thead>
<tr>
<th>Source images characteristics</th>
<th>Charlesbourg</th>
<th>Beauport</th>
<th>Montreal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scale</td>
<td>1/5000</td>
<td>1/8000</td>
<td>1/4000</td>
</tr>
<tr>
<td>Camera type</td>
<td>Digital</td>
<td>Film</td>
<td>Film</td>
</tr>
<tr>
<td>Focal length</td>
<td>120 mm</td>
<td>152 mm</td>
<td>306 mm</td>
</tr>
<tr>
<td>Pixel size</td>
<td>12 microns</td>
<td>21 microns</td>
<td>14 microns</td>
</tr>
</tbody>
</table>

The principle of MRAP is to define a library of objects consisting of several 3D parametric models and 2D geometric patterns, and eventually consisting of generalization algorithms and constraints. Each parametric model and geometric pattern of the MRAP corresponds to a single predefined abstraction level. All the parametric models and geometric patterns included in a MRAP are linked to each other. Knowing the parameters of the more detailed parametric model of the MRAP allows for inferring the implementation parameters of the other parametric models and geometric patterns included in the MRAP. When creating a new MRAP, the computation method of the coarse level parameters using the fine level parameters must be specified by a cartographer during an a priori step. Fig. 2 represents some MRAPs identified in our test sites (defined in Table 1, Section 4.1).

Four types of basic volumetric primitives can be used to create a 3D parametric model. They are illustrated in Fig. 1. When defining a geometry with such volumetric primitives (combined in a parametric model), we need to define its global parameters (i.e. implementation point and orientation in the horizontal plane) and its shape parameters. Shape parameters of the volumetric primitive number two are also represented ($w$: width, $l$: length, $\alpha$: roof slope, $h$: building height).

Fig. 2. MRAPs examples.
MRAP is an extension of the Multi-Scale Pattern concept, proposed in Cardenas (2004). The latter is restricted to 2D and tackles only CLG geometry extraction. Moreover, it involves neither constraints nor generalization algorithms. Adding those constraints and algorithms to the MRAP allows for supporting the SGO concept (currently under development in our research group) (Sabo, 2007). Linking the multi-scale pattern concept with the parametric model makes the MRAP usable in the FLG photogrammetric acquisition context.

Data acquisition with MRAP consists in implementing (i.e. setting the following parameters: anchor point, rotation, height, ...) the most detailed 3D parametric model of the MRAP. It is also possible to define groups of geographic objects that must fulfill the same generalization constraints (e.g.: the buildings belonging to the group must all be lined up during the generalization) or the same generalization rule (e.g.: the buildings belonging to the group must all be aggregated up during the generalization).

The example displayed in Fig. 3 shows how the MR acquisition procedure provides several abstraction levels. They range from a fine geometry per building to a coarse geometry for a group of aligned buildings. In this example, three building groups fulfilling the same aggregation rule have been built. More specifically, Fig. 3(d) illustrates how an aggregation rule, introduced during the capture, can be used to aggregate objects and thus produce more abstract levels. If we refer to the CityGML initiative, Fig. 3(b) corresponds to LOD2 and Fig. 3(c), (d) correspond to LOD1.

Acquisition results, involving the use of MRAPs 1, 4 and 5 (cf. Fig. 2), have been obtained using our manual prototype (Frédérique et al., 2008) on the Beauport test site (test site description is provided at the end of the paper, Table 1; Section 4.1). The human operator can select the relevant MRAP model using the corresponding tool selection designed in the prototype. MRAPs can be grouped during the acquisition process if they have to fulfill the same generalization constraint or rule. The manual MR acquisition prototype allows for determining the MRAP parameters from the selection of specific points in the 3D ground space (e.g. the two first points provide the length and the rotation angle of the MRAP). A human operator performs this selection in a stereoscopic view. One of the points must be located at the ground level. The other points must refer to specific vertices of the MRAP’s finest level. For example manual MR acquisition with MRAP number five (cf. Fig. 2) requires the selection of five points. This can involve up to seven mouse clicks if the corresponding tool must be selected and a building group must be defined. This number of clicks increases rapidly with the number of MRAP parameters. The first algorithm proposed in the next section aims to reduce the number of required points for the MRAP acquisition.

2.4. The limits of MRAP use

As mentioned before, we are aware that MRAPs are not relevant to all building geometries and that more flexible approaches must be considered to tackle complex building representations. Some building elements (e.g. roof window) are not always included at the finest level of the MRAP. It does not mean the MRAP cannot be used to extract the other building abstraction levels. More specifically the fact that there is no parametric model to describe the finest level does not imply that parametric models and geometric patterns cannot be used for coarse levels. The coarser the abstraction levels, the more relevant parametric models and geometric patterns are. Practically, it would consist in extracting a detailed geometry through an independent process (manual or semi-automatic) and
then linking this detailed geometry to a MRAP. The link between the different coarse levels and the parameters would be automatically computed through the MRAP. Regarding the finest level geometries, they could be defined using different geometric structures like Brep or volumetric primitives combined in CSG. Ideally, the CSG structures should be favored for database integrity optimization and to facilitate generalization operations (cf. 2.2.2).

To summarize, acquiring complex buildings in MR would involve two major components (1) the detailed geometry extraction, and (2) linking the detailed geometry and the MRAP and then inferring the MRAP parameters. The second algorithm proposed hereafter aims at supporting detailed geometries extraction and to facilitate the matching between detailed geometries and MRAPs.

3. Semi-automatic MR acquisition with orthogonal MRAPs

The following section deals with contexts where MRAPs are detailed enough to represent the finest geometry level of the buildings. The automatic definition of MRAP parameters (i.e. implantation parameters of the most detailed parametric model) follows the selection of the MRAP type and the building’s approximate location by the human operator. This issue corresponds to the determination of the parameters of the finest parametric model encapsulated in the MRAP. All identified MRAPs are orthogonal (i.e. they consist of primitives that are orthogonal to each other). This a priori knowledge about the volumetric primitives’ spatial relations will be useful for the extraction of their implementation parameters.

3.1. Works related to the parametric model parameters determination

Several research works have been dedicated to the determination of the parametric model parameters. Some of them are described hereafter. In fact, a priori knowledge regarding building geometries can easily be formalized and introduced using parametric models. A priori knowledge is a key element in the performance of building extraction algorithms (Baltsavias, 2004). Our global strategy, involving a top-down approach and the use of a priori knowledge, relies on this statement.

Approaches using parametric models can be classified in two categories. The first consists in automatically identifying the parametric model type as proposed in Haala and Brenner (1999), Suveg and Vosselman (2004), Lafarge et al. (2006) and Ortner et al. (2007). The second requires a human intervention in order to specify the parametric model such as in Gülch et al. (1999), Vosselman and Veldhuis (1999), Rottensteiner and Schulze (2003) and in Tseng and Wang (2003).

Works of the first category look for a full automation even if some authors also provide semi-automatic tools to post-process failure cases. Parametric models are then used to generate and evaluate hypotheses about the building geometries. Hypotheses are evaluated by comparing the hypothetical geometries with reference data sources consisting mostly of aerial images and DSM. DSM are more and more used since, as mentioned in Brenner (2005), information extracted from DSM simplifies automatic reconstruction issues as soon as this information is already in the object space. Furthermore, the recent progress of LIDAR sensors and aerial digital cameras in the last decade has improved the quality of the available DSM. Vector data can be used in addition to aerial images and DSM. In Haala and Brenner (1999) and Suveg and Vosselman (2004), building footprints are used jointly with a DSM. The proposed algorithms, using volumetric primitive combinations, rely on the segmentation of the building footprints. Gerke et al. (2001) and Vinson and Cohen (2002) proposed two automatic approaches to extract building footprints. Both strategies assume that building footprints can be described as rectangle combinations. Rectangles can be considered as parametric models restricted to the horizontal plane. These two strategies correspond to a generalization of the inertial moments based on the Maas (1999) method.

Semi-automatic approaches using parametric models, as proposed in Gülch et al. (1999), Vosselman and Veldhuis (1999), Rottensteiner and Schulze (2003) and in Tseng and Wang (2003), require a human to intervene at the beginning of the process. The purpose of this intervention is to identify the parametric model type and to define approximate parameters for the model. Accurate parameters are then computed automatically. The human operator can also combine parametric models to deal with buildings having a complex geometry. Parametric models are restricted to those having quadrilateral footprints, generally rectangles. Approximated parameters definition requires several clicks. For example, in Rottensteiner and Schulze (2003) three specific points are needed. Fine parameters are computed through optimization procedures that consider consistency between parametric model position and reference data. For example, in Rottensteiner and Schulze (2003) and in Tseng and Wang (2003) aerial image data is used as reference data and parametric model wireframes are
projected to edge images. Maas (1999) proposed a strategy quickly providing approximate parameters from a DSM region, a parametric model type and a unique approximated point. This strategy is based on inertial moments and can only be used with elementary parametric models (composed of a single volumetric primitive). Those cannot be used to combine primitives in CSG unless the regions of interest in the DSM, corresponding to each primitive, are specified earlier. Furthermore, as mentioned in Gerke et al. (2001), inertial moments are unstable if the shape of a building footprint is very similar to a square. The strategy proposed in Maas (1999) still has the major advantage of being quick and not demanding from the human intervention standpoint. These two aspects are particularly relevant in a semi-automatic context where response times must be short to keep the human operator from waiting, and where human intervention should be reduced as much as possible.

Most of the existing approaches, fully or semi automatic, use elementary parametric models (composed of only one volumetric primitive). They combine these elementary parametric models to describe complex geometries. Therefore, few algorithms are available to determine the approximate parameters of complex parametric models (composed of several volumetric primitives combined in a specific way) from a restricted number of approximated points. Strategies such as clustering or RANSAC (Fischler and Bolles, 1981) used for an elementary parametric model (Gulch et al., 1998) can hardly be used for complex parametric models as soon as the number of unknown parameters increases. This is particularly true in a semi-automatic context (i.e. needing real time processing). Existing algorithms computing accurate parameters from approximated values are still relevant insofar as each primitive optimization can be addressed independently.

3.2. Proposed algorithm for the orthogonal MRAP parameters determination

The speed of the automatic process is of great importance in our semi-automatic context. In fact, implementation parameters must be established on-the-fly. Several of the MRAP involve complex parametric models (involving several volumetric primitives). This complexity may imply many parameters to optimize and increase the automation difficulties. Moreover, to limit human intervention, we look for an initialization stage consisting in only introducing an approximate point. We restricted the spectrum of our investigation to the estimation of approximated parameters of a complex MRAP from an approximated point. In fact, as mentioned before, several contributions already tackled the accurate parameter determination from an approximated initialization.

3.2.1. Strategy for the determination of parameters: Extraction and selection of primitives

We have introduced a new approach using passive imagery and DSM to extract the parameters of orthogonal parametric models (the more detailed parametric model included in an orthogonal MRAP).

A DSM is automatically produced from stereoscopic couples by combining area-based matching techniques (correlation from object space) and optimal surface research in a graph (minimum cut research). Introduced by Roy (1999), this approach allows for a spatial homogeneity notion to be included during the DSM computation process. Thus, it uses a priori knowledge regarding the area continuity and allows for enhancing the robustness of the DSM extraction. The reader can refer to Pierrot-Deseilligny and Paparoditis (2006) to get a detailed description of this DSM production method. Fig. 4 is an example of an image source and its corresponding DSM generated using the proposed
strategy. This example refers to our second test site (Charlesbourg) described in Table 1, Section 4.1.

Our strategy to automatically extract the parametric model parameters from a single approximated point divides the optimization problem into several subproblems to solve it quickly. This strategy considers a parametric model as a particular volumetric primitive combination. The parametric model parameters’ research is then addressed by extracting the parameters of each volumetric primitive. This strategy involves three successive steps: (1) determining a Region Of Interest (ROI) in the ground space including the building to be extracted, (2) extracting the best volumetric primitives combination according to the parametric model, and (3) deducing the parametric model parameters from the parameters of each primitive and the connectivity relations between these primitives.

The ROI is extracted using a normalized DSM (nDSM). The nDSM is computed by subtracting the Digital Terrain Model from the DSM. The nDSM is segmented using a region-growing algorithm based on the height analysis (Weidner, 1997; Ameri, 2000). The height threshold is defined according to the a priori knowledge regarding the minimum height of the building. The DTM is deduced from the DSM through the application of a morphological filter (Weidner, 1997; Sternberg, 1983). The DSM, nDSM and ROI are described using a raster format. Examples of source image, nDSM, ROI and corresponding MRAP are shown in Fig. 5. Sub-step 2.1 consists of converting the ROI from a raster form to an orthogonal combination of rectangles. Again, the developed approach is considered as a contribution to this research field. The principle of the ROI conversion strategy is to generate a collection of axes that are subsequently used to generate rectangles through a dilatation process. An axis is defined by a point and a direction, and the dilatation process aims at defining the rectangle dimension around this point. Starting from the fact that thinning methods generate axes with directions that are strongly affected by the noise of the shape boundaries (Parker, 1997) and that the positions of the axes are relatively robust, we have proposed a mixed strategy to express the ROI as a rectangle combination. This strategy uses both the raster ROI and the edge images to express the ROI as an orthogonal combination of rectangles. The well-known Canny filter (Canny, 1986) is used to compute the edge images from the source images. We first extract rectangle axes (defined with a

Sub-step 2.1 consists of converting the ROI from a raster form to an orthogonal combination of rectangles. Again, the developed approach is considered as a contribution to this research field. The principle of the ROI conversion strategy is to generate a collection of axes that are subsequently used to generate rectangles through a dilatation process. An axis is defined by a point and a direction, and the dilatation process aims at defining the rectangle dimension around this point. Starting from the fact that thinning methods generate axes with directions that are strongly affected by the noise of the shape boundaries (Parker, 1997) and that the positions of the axes are relatively robust, we have proposed a mixed strategy to express the ROI as a rectangle combination. This strategy uses both the raster ROI and the edge images to express the ROI as an orthogonal combination of rectangles. The well-known Canny filter (Canny, 1986) is used to compute the edge images from the source images. We first extract rectangle axes (defined with a...
point and an orientation) and then rectangle dimensions around these axes (Half Forward Length (HFL), Half Backward Length (HBL), Half Right Width (HRW), and Half Left Width (HLW)). We use the edge images to extract the axes direction and the ROI skeleton to define the approximate location of the axes. The Zhang-Suen (Zhang and Fu, 1984) thinning algorithm is used to compute the skeleton from the raster ROI. Rectangles are then defined using these axes (to initialize the rectangles) and the raster ROI (to delineate the rectangles around the axis). Such a strategy is relevant since the building direction is accurate in the edge images while the ROI boundaries are noisy. The edge images cannot be used to improve the rectangle boundaries. In fact, it would be necessary to distinguish edges near the building boundaries from the edges belonging to the inner roof structure. Edge images will be used during the extrusion process to improve the boundary locations of the volumetric primitives.

3.2.2.1. Determination of the principal direction. The principal direction is defined between 0° and 90° and corresponds to volumetric primitive directions (primitives are orthogonal). The principal direction is deduced from the edge images using the statistical analysis of the edge directions included in an analysis area. The analysis area is deduced from the ROI through its dilation according to the DSM accuracy. In the analysis area, edge images are orthorectified using the average altitude in the ROI. This strategy is usable if the camera focal that is used for image acquisition is short enough to confine edges inside the analysis area. In fact, at the extrusion step, roof structure knowledge is available and it can be used to select relevant edges. The ROI conversion strategy is synthesized in the UML activity diagram of Fig. 7. Both the extraction of direction and the rectangle definition from the axis are described hereafter.

Fig. 6. UML activity diagram of the global process.
the Z-variation of the roof structure edges should not involve a delocalization of the building edges outside the analysis area. This strategy is not only very simple to implement but it is also very fast. It is based on the a priori knowledge that most of the edges are in the principal direction of the building (i.e. the building roof structures yielding to edges in the images are horizontal) and that the determination of the edge direction is robust.

3.2.2.2. Generation of the rectangles from the approximate axes. The axes are generated from both the skeleton and the orthogonal MRAP principal direction. For each line of the skeleton we create two axes defined by a point and an orientation. We use the middle point of the skeleton line to define the axes. The orientations of the two axes refer to the building principal direction and to the building principal direction plus 90°, respectively.

For each axis, a heuristic, illustrated in Fig. 8, which allows for the estimation of the largest oriented rectangle included mainly in the ROI, is used. The rectangle should include the axis point. A rectangle is defined by a point and five other parameters: orientation, Half Forward Length (HFL), Half Right Width (HRW), and Half Left Width (HLW).

The heuristic involves three main steps described hereafter and is illustrated in Fig. 8:

1. First estimation of HFL and HBL
   - Intersection of the axis and the ROI boundaries
   - Creation of segment S1 with the two shortest points in the forward and backward directions.

2. Determination of HRW and HLW
   - Statistical analysis (first quartile) of the distances between the points belonging to the ROI boundaries and their orthogonal projection on segment S1
   - Note: only orthogonal projections inside segment S1 are used.

3. Final adjustment of HLF and HBL
   - Creation of segment S2 from a point (axis point), a direction (perpendicular to S1), an Half Forward Length (corresponding to the previous HRW) and an Half Backward Length (corresponding to the previous HLW).
   - Statistical analysis (first quartile) of distances between the points belonging to the ROI boundaries and their orthogonal projection on segment S2.

Fig. 7. UML diagram activity describing ROI conversion algorithm.
3.2.2.3. Selection of the best orthogonal combination of rectangles. Finding the best orthogonal combination of rectangles comes after the expression of the ROI as an orthogonal combination of rectangles. This research is based on an a priori knowledge about the building footprint (e.g. L shape or T shape). This a priori knowledge, specific to the instance, is input through the identification of the MRAP type and through the approximated point input. The approximated point links the MRAP type to the ROI.

The best orthogonal combination of rectangles results from matching and evaluating the combination of rectangles with the parametric model footprint. Since graph structures are relevant when performing topological comparisons, the matching process is performed according to a graph-based strategy (research of the graph and subgraph isomorphism). Each rectangle is a graph node and each overlap between two rectangles is a graph edge. This definition of the graph edge depends on the relative positions of the axes of the two overlapped rectangles. The best subgraph maximizes a score function and must overlap the approximated point introduced by the human operator.

We have used the area generated by the rectangle’s graph footprint (logical OR) as the score function. Other score functions could be used. For example, in Ortner et al. (2007), authors check the consistency between the theoretical roof structure (defined as a volumetric primitive) and the DSM in order to evaluate the relevance of the solution. To identify the best subgraph we use a heuristic assuming that the best subgraph likely includes the biggest rectangle overlapping the approximate point. All the possible sub-graphs including this node are evaluated. If no solution is found, the second best rectangle using the approximate point is then used. The same principle is iterated until a solution is found. Fig. 9 illustrates the graph matching process used to compute the best combination of rectangles.

3.2.3. Extrusion and segmentation

The selected combination of rectangles describes the initial building footprint. The volumetric data is extracted during the extrusion step. It uses both preliminary extracted rectangles and knowledge about the volumetric type. This knowledge is attached to the rectangles during the graph-matching process.
(knowledge example regarding the roof structure: gable roof). On the one hand, extrusion aims at extracting inner roof structure. On the other hand, it aims at extracting building height. Additional elements, such as trees and neighboring buildings, can be included in the footprint. They are excluded during the segmentation steps.

Thus, each rectangle included in the selected sub-graph is extruded and segmented in order to generate a volumetric primitive. Extrusion and segmentation processes are specific to the volumetric type. Rectangles are extruded according to the volumetric primitive types. For each type of volumetric primitive, extrusion and segmentation processes exploit faces, defined in the 3D object space, that have been preliminary extracted. The faces extraction is performed in the same way for each volumetric primitive type. A score is attached to each face during its extraction.

Segmentation relies on topographic and radiometric discontinuities. In fact, the ROI analysis is not sufficient to delineate a volumetric primitive. A first segmentation, based on topographic discontinuities, is performed during the extrusion step. Then, a second segmentation, based on the image’s edge analysis, is carried out. Rectangle position in the best sub-graph is considered during the segmentation in order to maintain the selected graph consistent with the parametric model footprint. As mentioned before, the selected graph must display the same topological structure as the parametric model footprint (e.g. T shape) and must overlap the approximate point.

The extrusion and segmentation processes are described hereafter. Extruded roof boundaries are projected to the ground in order to compute the volumetric primitive height.

3.2.3.1. Extraction of faces. All the faces are extracted for each rectangle used to generate a volumetric primitive. A 3D-point cloud is computed from the DSM. For this purpose, we define a regular grid inside the rectangle to be extruded. The XY coordinates of the points are computed through the rectangle discretization and the Z coordinate is deduced from the DSM. The planes are computed from this point cloud using a RANSAC strategy followed by a least square compensation. The face creation consists in delimiting each plane with a polygon. The polygon is deduced from the oriented bounding box of the selected cloud points. The initial bounding box is enlarged using a buffer defined according to the DSM accuracy and the grid-sampling step. All the extracted planes through the RANSAC strategy are kept but a score is attached to each face depending on the face area and the number of points used to extract the plane. Fig. 10 illustrates the extrusion process for a gable roof volumetric primitive. This primitive refers to the horizontal part of the running example T shape.

RANSAC is a reliable approach that can be easily implemented to extract planes. However, inadequate settings or particular configurations can lead to method failure. The threshold criterion used to accept or reject a point according to its distance to the current plane must be set carefully. In fact, a threshold that is too small will lead to the extraction of a lot of vertical planes parallel to the sampling direction of the grid. A threshold that is too high will affect the accuracy of the extracted plane because outlier points will be used for the least square adjustment of the plane equation. Moreover, these points will be rejected during the next computation and they will not be involved in other plane computations. Consequently, an efficient extraction of the faces relies on the knowledge of the DSM accuracy.


3.2.3.2. Extrusion of flat roof and pent roof primitive and segmentation from a topographic standpoint.

Extrusion of flat roof and pent roof primitives (primitives 1 and 2, Fig. 1) needs to select the best face among all the extracted faces. This face selection depends on both the graph topological constraint (consistency between the parametric model and the overlap with the approximate point) and the score attached to the face. If the volumetric primitive to be extruded is a flat roof primitive (primitive 1, Fig. 1), the selection is restricted to horizontal faces. The selected face is used to segment the primitive footprint by keeping only the overlapped area between the face footprint and the initial rectangle. This segmentation is performed according to topographic discontinuity criterion. A DSM section corresponding to several flat roof primitives is shown in Fig. 11. The rectangle displayed on Fig. 11(c) represents the initial rectangle. Rectangle produced by the segmentation is shown on Fig. 11(e). This example is extracted from our Beauport test site.

Volumetric primitive parameters are then deduced. Implementation point, orientation, width and length are deduced from the rectangle. The selected face is then used to define building height (vertical distance between face and DSM at the implementation point) and roof slope (face slope).

3.2.3.3. Extrusion of a gable roof primitive and segmentation from a topographic standpoint.

Extrusion of a gable roof primitive (primitive 3, Fig. 1) needs to select the best two faces that intersect each other in a consistent way according to the volumetric primitive shape. In fact, the intersection of the two faces must correspond to a peak and not to a thalweg or a simple slope discontinuity. This face selection depends also on both the graph topological constraint (consistency between the parametric model and the overlap with the approximate point) and the score attached to the faces. The face selection is performed through the identification of the best peak 3D edge. A 3D edge is computed for each intersecting faces pair. These 3D edges are then classified into two categories: peak and discontinuity. A 3D edge is a peak if the Z values of its two end points are higher than the Z values of the two faces’ centroids. All other edges belong to the discontinuity class.

A score function, using the score attached to the two faces and the 3D edge length, is then used to select the best peak edge. Only peak edges displaying the same direction as the principal direction of the rectangle are considered. The segmentation of the footprint, based on topographic considerations, follows the same principle as the flat roof case, where we keep only the overlap area between the oriented bounding box of the two faces footprint and the initial rectangle.

The selected peak edge and the segmented footprint are then used to deduce the volumetric primitive parameters. The rectangle is used in a similar way as for flat roof primitive. The 3D edge is used to define roof structure parameters. The peak height is deduced from the 3D edge middle point height and the peak eccentricity is deduced from 3D edge distance to rectangle boundaries.

3.2.3.4. Extrusion of a hip roof primitive and segmentation from a topographic standpoint.

A hip roof primitive is processed as a particular gable roof primitive with an additional parameter. This parameter defines the shift between the beginning of the peak and the primitive footprint. It is computed by adjusting the peak edge with the normal faces (the orientation of the normal vector of the face in the XY plane is the same as the peak edge).

3.2.3.5. Radiometric based segmentation.

A volumetric primitive is delimited not only by its 3D shape (topography discontinuity in the DSM), but also by edges...
and/or changes in the spectral properties. According to this definition, there are seven volumetric primitives in Fig. 12 (considering small details in the roof boundaries as negligible). In order to acquire one of these seven connected buildings, we need to consider not only the topographic discontinuities but also the radiometric discontinuities.

Fig. 13(a) illustrates the face extracted in the context of Fig. 12 image when the radiometric discontinuity is not considered. In order to extract the volumetric boundaries, additional processing is required after the extrusion and topographical segmentation. This processing consists in segmenting the volumetric boundaries using the edge images. For this purpose, we look for 3D edges in the object space that delineate the primitive footprint as a rectangle. More specifically we look for 3D edges that are (1) oriented according to the principal direction, (2) on the roof 3D faces (i.e. on the extracted and selected faces when defining the primitive roof structure at the previous step).

Using knowledge about the roof structure type, we generate 3D edge candidates (oriented in the principal direction and on the roof 3D faces) by setting their length to the initial rectangle dimensions and by discretizing the candidate space according to the DSM ground resolution. Fig. 13(b) and (c) illustrate the 3D edge candidates generated in the context of Fig. 12 image. The relevance of the 3D edge candidates is then assessed using the edge image. The 3D edge is projected on the edge image. The discrepancy between its position and the position of the detected 2D edges in its vicinity provides a measure of the 3D edge relevance. The number of edge pixel in this neighborhood is directly used as a measure function. The neighborhood is defined according to the DSM accuracy and the edge image resolution.

The radiometric based segmentation then consists in finding the 3D edges with a measure higher than a fixed threshold. The research is performed according to four directions (principal direction modulo \( \pi \)) around the approximate point, starting from the nearest 3D edge to the furthest. For each direction, the search stops as soon as a measure value higher than the threshold is found. If the approximated point is not inside the volumetric primitive (when the parametric model consists of several primitives), a point defined contextually is used. If the roof type is gable or hip we use a point on the peak axis (the peak was selected during the previous step). If the roof type is flat or pent and if the footprint is a Tshape we use a point on the medial axis of the initial rectangle, etc.

4. Experimentation

4.1. Test sites

The processes used to automate the determination of orthogonal MRAP parameters have been carried out on three test sites in the province of Quebec. The first two sites are near Quebec City (Beauport and Charlesbourg) and the third is on the Island of Montreal. These three test sites consist of medium-density residential areas (relatively small buildings, less than 4 floors). This kind of urbanization is a typical North-America suburb configuration. The orthogonal MRAPs are relevant in such a context. The characteristics of the aerial images we used are described in Table 1. DSMs were generated with a 25 cm ground resolution because we were only looking for approximate determination of the MRAP parameters. Furthermore, having the same
ground resolution for all the DSMs will facilitate the comparison of the various results.

4.2. Results

We compared the parametric models that were automatically implemented with the manual implementation (using stereoscopic images) in order to validate the automatic determination of parameters. Results are considered valid when: (1) the difference between the automatically and the manually positioned edges are smaller than three times the DSM resolution, and (2) the angular difference between the two principal directions (i.e. automatically and manually extracted) is smaller than 3°. These criteria were based on our review of works related to fine parameter determination from the manual introduction of approximated parameters (ex. results in Gülch et al., 1999).

Our extraction strategy of the initial MRAP's parameters relies essentially on the DSM. Thus, the resulting accuracy is intrinsically dependent on the DSM accuracy. The resulting accuracy could be improved with a better DSM. The implemented prototype involves computation times smaller than 1 second per volumetric primitive using a laptop with the following characteristics: Intel Centrino 1.7 GHz, 512 Mo RAM. Some of the processes (e.g. face extraction, edge direction computation) should be computed during a preliminary step in order to decrease the on-the-fly time processing. The table below synthesizes the obtained results. Fig. 14 illustrates the result obtained on the Charlesbourg test site.

4.3. Analysis

The implemented processes dedicated to the extraction of MRAPs’ parameters have been assessed on three
Table 2
Results for the three test sites

<table>
<thead>
<tr>
<th>Test Site</th>
<th>Total number of tested buildings</th>
<th>Failure during initial footprint extraction</th>
<th>Failure during extrusion and segmentation</th>
<th>Additional knowledge about principal direction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Charlesbourg</td>
<td>107</td>
<td>ROI detection 1 0 0 4 5.3 0 0 1.3</td>
<td>DSM resolution is not enough to extract element that are differentiable in images 4 2 0 1 0 0 0 0.3</td>
<td>Information about principal direction available through groups constitution 11 4 0.7 0 0 0 0 0.3</td>
</tr>
<tr>
<td>Beauport</td>
<td>76</td>
<td>Principal direction detection 4 3.7 0 0 10 0 7.7 4.5</td>
<td>Noise in the DSM surfaces 5 0 0 10 13 40 30.7 16.0</td>
<td>Information about principal direction can be used to bypass failure in principal direction extraction 10 43 40.2 68 89.5 100 76.9 67.4</td>
</tr>
<tr>
<td>Montreal</td>
<td>130</td>
<td>Rectangle extraction according to parametric model (graph matching) 3 2 1.9 3 4 4 1.5 2.9</td>
<td>Details elements (ex: windows) 6 2 1.9 0 0 2 1 1.3</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>100</td>
<td></td>
<td>Abusive segmentation 7 0 0 0 0 2 1.5 0.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Failure during extrusion and segmentation</td>
<td>Information about principal direction available through groups constitution 11 4 0.7 0 0 0 0 0.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Information about principal direction can be used to bypass failure in principal direction extraction 10 43 40.2 68 89.5 100 76.9 67.4</td>
<td></td>
</tr>
</tbody>
</table>

Some failures can arise during the initial footprint extraction (ROI extraction, conversion to a rectangle graph structure...) and during the extrusion and segmentation step. The ROI extraction is very robust (a 99% success rate) even if the DSM quality is relatively poor. The principal errors arise when the nDSM quality does not allow for building detection during the ROI extraction. The suggested strategy, which is dedicated to the extraction of the building principal directions and the description of the ROI as a rectangle graph, leads to a failure rate of less than 10% for the three test sites. The principal cause of failure during the direction extraction is tree branches. Using edge ortho-images, the automatic determination of the principal direction could be improved through a 3D edges analysis instead of a 2D edges analysis. The improvement would consist in extracting the 3D edges with a feature-based matching strategy and then in analyzing only the horizontal 3D edge directions. Furthermore, it should be underlined that a priori knowledge about the principal direction of the building to be extracted was available in all the failure cases (lines 10 and 11 in Table 2). We indeed counted the number of occurrences when this kind of a priori knowledge was available. The purpose was to evaluate whether this available information could be useful from an automation standpoint. This information has been provided through the CLG specification when a group of aligned buildings were built. Taking advantage of this information when available is a prospect that should be investigated (e.g. cancellation of the edge analysis or exploitation of this information during the statistical analysis of the edge directions). The matching process between the available rectangles and the geometric pattern footprint displays a rough 3% failure rate. These failure cases occur when the DSM is too noisy and, consequently, when the ROI boundaries are too poor. The footprint extraction, particularly the rectangle graph computation, is still very efficient.

The initial footprint extraction processes are fast (first part of the Table 2), relatively robust and can take advantage of the available a priori knowledge about the building directions. The extrusion and topographical based segmentation are only based on the DSM exploitation. Our test sites are similar from an architectural standpoint (simple building structure compatible with parametric model concept). Therefore the performance differences between these test sites are attributable to DSM quality differences. In fact, noisy surface definition in DSM leads to extracting wrong faces and then leads to extrusion failure. Fig. 15 represents the shaded view of the three DSM and illustrates how the Beauport DSM and the Montreal DSM are noisier than Charlesbourg DSM. Several factors have an impact on DSM quality (image resolution, b/h ratio..) but we think the most
The deterministic factor in our case is likely the type of camera used (film vs. digital) that, as described in Paparoditis et al. (2006), changes drastically the signal to noise ratio and then the correlation success.

The semi-automated approach implemented with MRAP decreases up to 5 times the human intervention required when performing a manual MRAP implantation. The proposed algorithm could be easily combined with other strategy used to determine accurate parameters like those proposed in Tseng and Wang (2003) or in Rottensteiner (2001).

5. Outlook on buildings footprint extraction

5.1. Goal and strategy

This section presents some preliminary results regarding the extraction of complex detailed geometries based on the strategy described in Section 2.4. We used the Amiens city dataset provided by IGN and the workgroup 8 of the ISPRS commission III (http://isprs.ign.fr/) to facilitate the comparison of our results with other approaches. The algorithm we developed consists in extracting the building footprints while expressing them as an unconstrained combination of rectangles (rectangles are not necessary orthogonal to each other). The purpose of such an approach is (1) to support MR acquisition when MRAPs are not orthogonal, and (2) to extract the fine geometries and link them with a MRAP when MRAPs are not detailed enough to extract the finest level geometries. In fact, preliminary footprint extraction could be combined with other algorithms requiring geometric footprint description. It could, for example, be combined either with the strategies proposed in Haala and Brenner (1999) or in Suveg and Vosselman (2004), which provide a CSG description of the building, or with the strategy proposed in Jibrini et al. (2000), which is more generic from the roof structures standpoint.

5.2. Building footprints algorithm

The proposed algorithm uses aerial imagery and DSM and consists of two successive steps: the footprint expression as a graph of rectangles and the simplification of this graph. This algorithm assumes that the roof structure consists mainly of 3D edges aligned in accordance with the rectangle principal direction.

The initial graph extraction is similar to the orthogonal MRAP extraction. However, we do not consider that rectangles must be orthogonal to each other anymore. Indeed, a ROI is extracted from nDSM and lines corresponding to the ROI skeleton are used to create rectangle axes. The directions are computed independently for each axis. Only the edges inside the ROI and close to the axis point are used for the statistical analysis whereas, in the orthogonal MRAPs configuration, all the edges in the ROI were used.

The final building footprint is then inferred by simplifying the graph to decrease the number of rectangles while preserving satisfying geometry fidelity. We developed an algorithm to perform this simplification. The basic idea is to delete the rectangles that do not contribute enough to the footprint description. We used the covered area as a contributing criterion. The algorithm first selects the rectangle providing the best contribution and deletes all the rectangles that do not contribute enough. The second best rectangle is then searched following the same principle and irrelevant rectangles are deleted according to the two first selected rectangles. This procedure is iterated while the number of remaining rectangles keeps on decreasing. This algorithm is summarized in the UML activity diagram presented in Fig. 16. Fig. 17 illustrates the results obtained for the two main steps of the algorithm. The images consist of a group of connected buildings.

Fig. 18 corresponds to the results obtained for two other groups of buildings. The processing time per building group was about twenty seconds. We can visually notice that results are really close to a manual acquisition based on DSM even if some rectangle connections are not perfect. Lafarge et al. (2006) observed similar problems on the same test site when they used parametric models to automatically extract building geometries with a marked point process. The
authors also proposed an algorithm to correct these artifacts.

6. Conclusion

The multi-representation acquisition specifications have been described in this paper. The results obtained using our prototypes have been presented in order to illustrate the MR acquisition process and results. We have introduced two new algorithms. The first one can be used to extract the parametric model parameters with a simple one-click initialization. This new approach is relevant to elementary parametric model (with only one volumetric primitive) and to more complex parametric models (ex: T or L shapes) and then can be used to support MR acquisition with MRAP. The comparison
of manual MR acquisition using MRAP with the acquisition using the proposed new algorithm has shown a 5 times reduction of the human intervention in the process. The current performances are encouraging since the human intervention is lessened to a single click per building and a response time around one second per building. The short response time is particularly relevant in our semi-automatic context. The analysis of the results has allowed us to underline the limitations of our approach and to propose some recommendations. In addition, we proposed a second new algorithm aiming to extract the building footprints and to express them as an unconstrained combination of rectangles. The preliminary results are as also very promising as this second algorithm also offers a very quick response times.

Acknowledgments

The authors wish to thank for its support the Industrial research chair in geospatial database for decision, support financed by the Natural Sciences and Engineering Research Council of Canada (NSERC), the Université Laval, Hydro-Québec, Research and Development Defense Canada, Natural Resources Canada, Ministère des Transports du Quebec, KHEOPS Technologies, Intélec Géomatique, Syntell, Holonics and DVP-GS.

The authors wish to thank for its NSERC Industrial Postgraduate Scholarship the Natural Sciences and Engineering Research Council of Canada (NSERC). The authors would also like to thank DVP-GS’s technological and financial support, as well as for providing high resolution images and vector data.

References


