Data integration for 3D modeling

Intégration des données pour la modélisation 3D

Rafika HAJJI

Department of Geodesy and surveying Engineering, Institute of Agronomy and Veterinary Medicine, BP 6202 Rabat-Institute, Rabat, Morocco

r.hajji@iav.ac.ma

Abstract

Nowadays, 3D city models are getting more popular day by day. Achievements in the 3D domain are notable. However, 3D modeling still faces some impediments to be generalized. New solutions such as linking 2D and 3D worlds through 2D/3D data integration should be investigated.

This paper presents an overview of some realizations that have marked out the 3D modeling field and highlights some related questions that are still in the research. It presents and argues the importance of bridging the 2D and 3D worlds as a promising solution to upgrade to integrated 3D solutions in the future. This idea is supported by a case study, to demonstrate how 2D/2.5D data collected from different sources can be integrated and reengineered to match the specifications of a 3D building model compatible with the CityGML standard.

Résumé

De nos jours, les modèles 3D sont de plus en plus populaires. Les réalisations dans le domaine de la 3D sont notables. Cependant, la modélisation 3D est toujours confrontée à certains obstacles qui entravent sa généralisation. De nouvelles solutions telles que la liaison des domaines 2D et 3D à travers l'intégration de données 2D / 3D doivent être investiguées.

L'article présente un aperçu de certaines réalisations qui ont marqué le domaine de la modélisation 3D et souligne certaines questions relatives qui sont encore dans le champ de la recherche. L'article présente et souligne l'importance de relier les mondes 2D et 3D en tant que solution prometteuse pour la migration vers des solutions 3D intégrées dans le futur. Cette idée est soutenue par une étude de cas qui démontre comment les données 2D/2.5D collectées auprès de différentes sources peuvent être intégrées selon les spécifications d'un modèle 3D de bâtiments compatible avec le standard CityGML.

Keywords

3D modeling, data integration, CityGML

Mots-clés

Modélisation 3D, intégration des données, *CityGML*

1. Introduction

The reality is tridimensional in nature. Paradoxically, its representation has been for long decades dominated by the metaphor of the paper map. The third dimension is often handled, but only by considering elevation as a function of location, a solution often termed "2.5D" (Goodchild, 2010). 2.5D representations would be sufficient for some applications like urban planning, but cannot efficiently deal with representing multi-level structures in applications requiring higher-dimensional models, such as vertical cadaster. Likewise, representation of indoor and outdoor objects in 3D models is of interest for several situations such as the evacuation of buildings in case of extraordinary circumstances (disastrous accidents, terrorist attacks, etc.) (Atila, Karas et al., 2013). In such applications, a veritable 3D model considering internal/external and eventually multi-level structures is required.

Nowadays, spatial data is generalized for worldwide and shared by an extended number of users. 2D/3D city models are now accessible over Internet, allowing several kinds of spatial services, such as navigation and simulation. Now, technologies make it easier to produce and have access to 3D information. Anyone with Internet access and mobile devices (smartphones, GPS, cameras …) has the ability to produce 2D/3D data voluntarily and make it available to a large community of users. An example of these forms of data crowdsourcing is the OpenStreetMap-3D platform. Crowd sourced data is generally optimized for exploration and visualization purposes, but cannot deal with a large spectrum of applications requiring geometric and semantic accurate 3D data. So, 3D reference models are needed to support "intelligent" applications and 3D spatial analysis.

Although 3D information is well established in many government agencies and private organizations that make significant 3D development, agencies with smaller budgets, especially in the local government, are generally the least expected to invest in 3D data production. Their eventual 3D achievements remain mediocre. These organizations are challenged by adopting a collaborative approach in 3D modeling and design by sharing cost and experiences in order to make 3D available, structured and formalized enough to be shared Consistently.

Organizations have spent too much time and money to build 2D topographic inventories with rich geometric and semantic information. Capitalization and reuse of this information, in order to inherit knowledge associated to 2D data, is a promising solution to reach basic requirements of a 3D model. This is a way to link 2D and 3D worlds towards future integrated 3D solutions.

This paper addresses the issue about 3D modeling of buildings. Firstly, we try to capture some important realizations about 3D modeling in terms of achievements, limits and opportunities (Section 2). Then, we address the issue about integrating 2D and 3D worlds (Section3). In section 4, some conceptual and technical issues about the design of a 3D model are discussed.

Finally, we present a case study about 2D/2.5D data integration to produce a 3D model compatible with the CityGML standard (Section4).

2. 3D achievements, limits and opportunities

The demand for modeling and handling 3D and 4D data sets has been rapidly growing during the last decades (Breunig and Zlatanova, 2011). Nowadays, 3D models are developed within various domains and for different purposes (Biljecki et al. 2015). 3D technological achievements are notable. In the present section, we analyze them with regards to three aspects: 3D in Computer Aided Design (CAD) and Geographic Information Systems (GIS), 3D formats and 3D scaling. At the end of the section, we capture the importance of linking the 2D and 3D worlds, as a challenging solution for 3D modeling.

2.1 3D in CAD & GIS

Nowadays, GIS and CAD systems are widely used in 3D modeling. The most fundamental difference is that GIS models the world as it exists, whereas CAD models artifacts yet to be produced (Ibraheem, Hassan et al, 2012). Certainly, CAD solutions provide high capabilities in 3D detailed urban modeling (Robles-Ortega, Ortega et al., 2012). But, still lacks the geospatial and topological properties that are the crucial characteristics of 3D GIS. Even if the difference between CAD and GIS systems is becoming blurred with the internal attributes and database linkages enhancing CAD's capabilities, dissimilarities that existed between GIS and CAD prevent 3D GIS from being able to utilize the existing sources of 3D data, mostly based on CAD software (Karimi and Akinci, 2010).

A 3D GIS is the appropriate system for storing, manipulating and visualizing spatial 3D data in urban environments (Robles-Ortega et al, 2012). It aims to provide basic functions (data capture, data structure, data manipulation, data analysis, and data presentation) similar to those available in 2D GIS (Karimi and Akinci, 2010). A 3D GIS must be able not only to store and visualize the objects in a 3D environment, but also to handle, manage and analyze them. Therefore, adding the third dimension cannot be considered as a simple extension of 2D GIS because additional structures and modules must be created to support the new functions (Robles-Ortega et al, 2012). In contrast to 2D GIS, 3D GIS interact with more information and require more advanced techniques to reconstruct a 3D representation of each distinct object. Despite the progress made in 3D technologies, most available solutions are essentially 2.5D and 3D functions which are currently incomplete. Two-dimensional spatial analysis cannot simply be applied to 3D spatial data sets for the purpose of 3D spatial analysis. 3D GIS should be able to perform volumetric modeling and reasoning, and promote a better understanding of the natural phenomenon (Karimi and Akinci, 2010).

2.2 3D formats

To enable data usability between heterogeneous environments, standardization exchange methods comprising both spatial and semantic information are required.

GML (Geography Markup Language), created by the Open Geospatial Consortium (OGC) is a typical example for data exchange. GML is an XML structure that provides a standardized encoding of types defined in the conceptual models specified by the ISO 19100 series of International Standards. Based on ISO 19107 standard, GML provides classes for 0D to 3D geometric primitives, 1D-3D composite geometries (e.g. CompositeSurface), and 0D-3D geometry aggregates (e.g. MultiSurface or MultiSolid);

CityGML is the standard for representing 3D real-world information. It is designed as an open data model and XML-based format for the storage and exchange of virtual 3D city models. CityGML is an object oriented model that allows semantics, geometry, topology and appearance characteristics to be stored per each object. CityGML defines five different levels of details. The most prominent are the Level Of Details (LOD) of buildings. According to CityGML, LOD0 is a 2D map where buildings are represented with their footprints. The superior levels are regarded as 3D models with different granularities. For example, in LOD1, buildings are represented as block models with flat roofs; buildings in LOD2 are block models with detailed roofs, buildings in LOD3 are represented with detailed external architecture, and LOD4 describes building interiors (Fig. 1). As an improved LOD specification of 3D building models, Biljecki et al. (2016) proposed a better definition of LODs and their refinement into a set of 16 LODs focused on the grade of the exterior geometry of buildings, which provide a stricter specification and allow less modelling freedom.

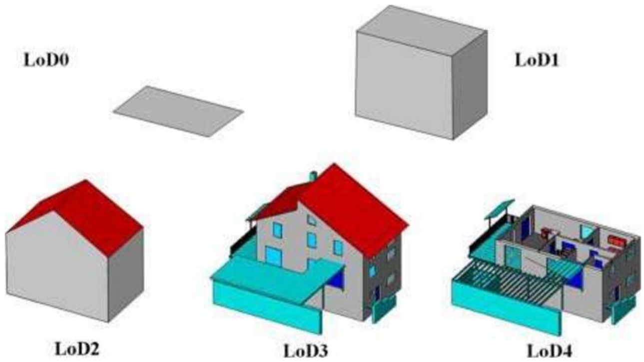

Fig. 1. CityGML LODs for buildings (Akmalia et al, 2014)

Other 3D formats exist. Each of them has its strengths and weaknesses. According to geometric criteria, Drawing eXchange Format (DXF), Virtual Reality Modeling Language (VRML), X3D, Collaborative Design Activity (COLLADA) and Industry Foundation Classes (IFC) support a large variety of geometries. VRML, X3D and COLLADA are the most advanced in supporting realistic textures. While standards such as Shp, IFC and CityGML have a very good support for semantics, CityGML scores relatively good on all criteria (Zlatanova, Stoter et al., 2012). (Gröger and Plümer, 2012) argue that there is no standardized rich semantic model for urban objects as CityGML. In fact, compared to other 3D formats, CityGML proposes a rich semantic and a multi-level representation which is very relevant in spatial applications and can be used as a source or target for generalizations (Gröger and Plümer, 2012).

Besides CityGML, IFC (Industry Foundation Class) is considered as another relevant milestone of 3D modeling of buildings. IFC provides detailed information about buildings, constructions and utilities which can be compared to CityGML LOD4 (Zlatanova et al, 2012) (See Fig. 2). So, integration of IFC and CityGML via a common exchange format is beneficial for enriching representation of urban space in different levels of details. (El- Mekawy et al, 2011); (Isikdag and Zlatanova, 2009); (El-Mekawy and Östman, 2010); (Benner et al, 2010) are some investigations dealing with this challenging issue.

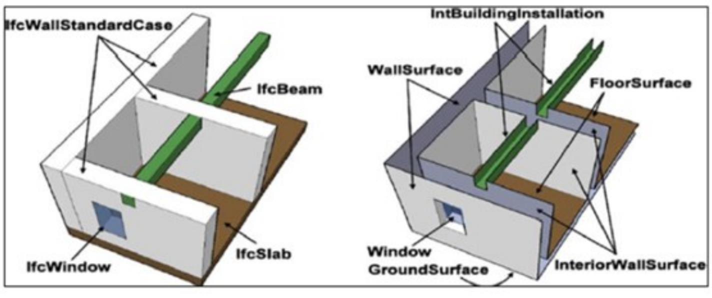

Fig. 2. Semantic representations of a building according IFC (left) and CityGML (right) (Nagel et al, 2009)

2.3 3D scaling

Considering that several 3D applications may share a unique 3D reference database, dealing with multi representations is a fundamental issue in 3D modeling. As it is defined in the CityGML specifications, the concept of level of detail allows representing objects with regard to different degrees of resolution which can be linked to the concept of scale (Van den Brink, Stoter et al., 2013). Each level of detail is characterized by differing accuracies (described as standard deviation of the absolute 3D point coordinates) and minimal dimensions of features (Gröger, Kolbe et al, 2008). Objects become more detailed with increasing LOD, regarding both geometry and thematic differentiation (Emgard and Zlatanova, 2007). Nevertheless, the accuracy and the description of how rich would be a representation of a building in each level of detail, as described in CityGML, might be confusing. For example, a building with internal description compared to LOD4, might be classified in a lower LOD because of a lack of accuracy which causes information loss. In this context, (Lowner et al., 2013) proposed a new Level of Detail concept that involves the separation of a Geometrical LOD (GLOD) from that of a Semantical LOD (SLOD). Furthermore, the consistent management of multiple representations is required to allow transformations between levels of details (Breunig and Zlatanova, 2011). In this context, Deng and Cheng (2015) proposed a methodology framework for automatic transformation of different LODs in CityGML.

Certainly, 3D achievements are notable. However, 3D modeling still faces some impediments to be generalized. This is due to lack of knowledge of 3D domain and lack of a generic approach to handle 3D Geo information (Stoter, Van dan Brink et al., 2011). Furthermore, 3D city modeling is a complex task due to the great amount of different urban entities to be represented (Robles-Ortega, Ortega et al., 2012) as well as the complexity of spatial analyzes. Therefore implementing a detailed 3D reference model would be a challenging solution for several organizations.

3. Bridging 2D and 3D worlds

In traditional GIS, objects have been modeled according two independent systems: DTM (Digital Terrain Model) and 2D GIS. Until GIS supports only the visualization of Digital Elevation Models (DEM), information derived from DTM must be converted into a thematic layer that GIS can recognize, before any spatial analysis can be carried out, resulting in data redundancy and inconsistency (Rahman and Pilouk, 2008).To overcome this limitation, one of the important directions for 3D geo-modeling concerns the integration of 2D and 3D worlds. Such a solution might avoid developing separate 2D and 3D solutions in the future.

Nevertheless, a principal limitation is that 2D GIS data models use a layer-based approach that prevents their direct application and extension to the 3D world. Integrating 2D and 3D data should consider which features in 2D maps to be taken in the 3D model, and which resolution features have to be represented as well as distinguish between domain specific 3D models (Breunig and Zlatanova, 2011).

To reach a future integrated 3D solution, basic objects of 3D models can be constructed by extending those of 2D topographic maps such as roads, buildings, etc. Therefore, 2D spatial databases can be reengineered and extended to build reference 3D models. This is a way to preserve semantic information generally rich for existing data and to reuse update processes, generally well established in 2D. Other motivations concern reducing the cost of data acquisition and assuring consistencies between 2D and 3D data.

In the next section, we present some conceptual and technical issues to deal with the design and the populating of a 3D reference model.

4. The 3D model: conceptual & technical issues

4.1 The design of a 3D model

A model is designed by an abstraction of reality in order to make it understandable to the users. It consists of a generic structure that can be populated with instances and includes classes, attributes, relationships, constraints and operations (Stoter, 2004). The consistency of a 3D model is more difficult to specify since user needs are diverse and poorly defined, in terms of object type representation and their geometric and semantic accuracy. In general, precision and detail levels are application-dependent. Additionally, a reference 3D object such as buildings must be assigned a standard semantic independently of organizational specificities.

As As previously stated, a 3D model should be interoperable. CityGML is seen as a reference 3D standard in order to ensure interoperability and to match the recommendations of the INSPIRE directive (Infrastructure of Spatial Information in Europe). Several solutions can be considered for representing 3D buildings according to four levels of detail of CityGML: a same LOD for all buildings, different LODs within the same data set (some buildings LOD1, others LOD2, LOD3 or LOD4) or make multiple representations in the same data set for the same building. The application schema "Buildings3D" of the INSPIRE directive is flexible. It only requires that at least one of the geometries (LOD1, 2, 3 or 4 shown in solid or multi area) is to be supplied through the constraint "MandatoryGeometry" (Fig. 3)

A 3D model should be based on a progressive approach. As a preliminary version, the 3D model would be generic and designed according to some common specifications based on standards (CityGML and INSPIRE directive). Data integration and reengineering will be engaged to populate the model by geometric instances and basic semantics. To be enriched in the future, data acquisition sources are to be defined according common specifications of geometric and semantic accuracy. As a prospective data source, crowdsource data could be integrated to the model by adopting a quality control process.

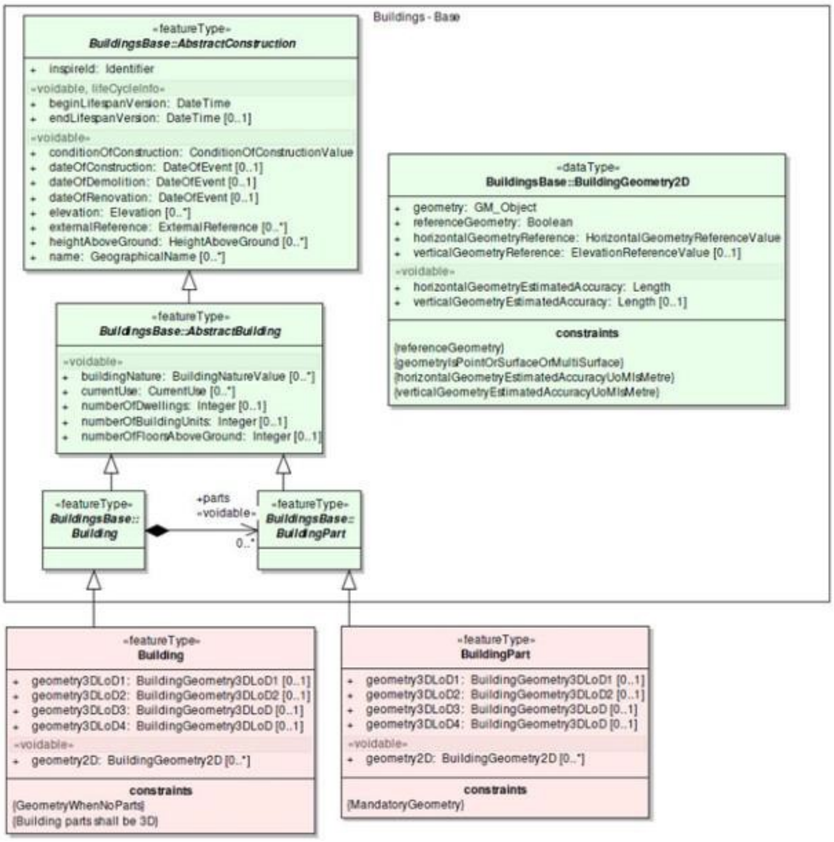

Fig. 3. The Building 3D profile (INSPIRE)

4.2 Populating the 3D model

Until the third dimension is recognized as a fundamental need, potential producers are likely to develop 3D geospatial databases. Nevertheless, acquiring detailed 3D models is not trivial. Knowing that each audience has a distinct set of 3D data needs, a major concern is how to establish one integrated solution to reconcile divergent and not yet well identified requirements for 3D information. We believe that defining and implementing a complete 3D reference model would be challenging, money and time consuming. A step-wise process may be more appropriate to reach basic requirements which can be enriched and completed since 3D reference geographic information is available.

Advanced data acquisition technologies such as 3D laser scanning and Unmanned Aerial Vehicle (UAV) systems have made 3D data acquisition more convenient and accessible for many applications. However, our reflection about 3D collaborative modeling as a step wise process relies on reuse and integration of 2D/2.5D data. For decades, 2D topographic inventories have been developed with rich geometric and semantic information. Therefore, capitalization and reuse of this information, in order to inherit knowledge associated to 2D data, is a promising solution to reach basic requirements of a 3D model. This is a way to link 2D and 3D worlds towards future integrated 3D solutions. In other words, a basic 3D model can be established from existing multi source 2D/2.5D data through a process of information integration and data reengineering. In fact, data integration aims to share and capitalize information from heterogeneous data sources in order to reuse it beyond the context of its creation. While data reengineering is required to match the specifications of the target 3D model.

Based on the assumption that basic objects of a 3D model are constructed by extending those of 2D topographic maps such as roads, buildings, etc. 2D models can be exploited as a building block to establish basic 3D models. Two-dimensional data are transformed into 3D objects, and new entities are created as well as the required topology to provide connectivity (Robles-Ortega et al.2012). In 3D Pilot project, Van den Brink et al (2013) have outlined four technical reasons to preserve information from existing 2D models: to maintain connection to applications which justify the use of 3D, to inherit semantics generally rich for existing data, to possibly generate automatically 3D models from existing data sets and to reuse update processes, generally well established in 2D. Other motivations concern reducing the cost of data acquisition, assuring consistencies between 2D and 3D data and gradually evolve to future integrated 3D solutions.

Finally, we can state that 2D/2.5D data integration is an essential step to reach a basic 3D reference model. Section 5 of this paper addresses this challenging issue through a case study which deals with 2D/2.5D data integration and reengineering in order to match specifications of a building 3D model compatible with CityGML. We have used data from Belgium but the methodology can be easily transferable to other contexts where 3D data is missing.

5. Data integration for 3D modeling

Existing 2D spatial inventories are generally established to respond to intra organizational needs and thus designed according different specifications. Furthermore, data representation of the same geographical area is usually dispersed and heterogeneous because of its “design autonomy”. Thereby, sharing and capitalizing information from such data sources require a process of information integration.

Databases integration is conducted through preintegration, correspondences investigation (data matching) and integration (Fig. 4)

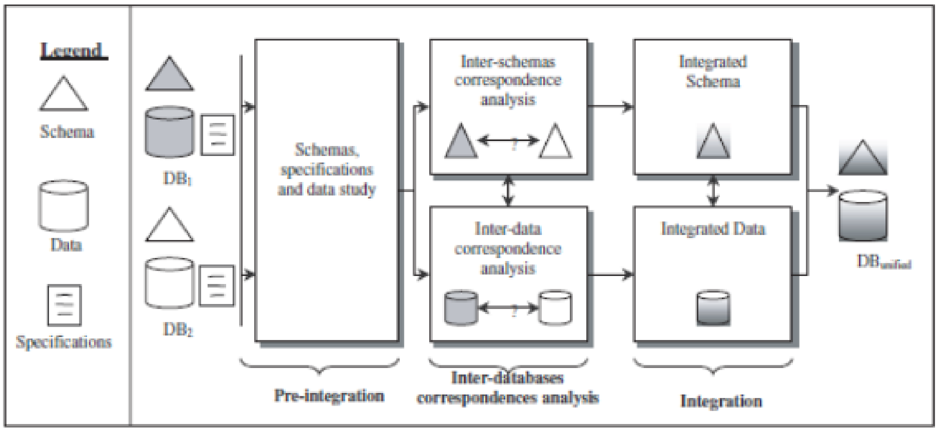

Fig. 4. Main steps of spatial databases integration process (Sheeren et al, 2004)

In this section, we present some results of a process of data integration and reengineering to reach a basic 3D model based on CityGML standard. To carry out the experiment, we have used 3 data sets of Liege City in Belgium: a cadastral database, a Geodatabase file from NGI (National Geographic Institute) database and the building layer of Walloon region topographic inventory (PICC).

We have fixed two main objectives for databases integration. The first one consists of extracting geometric updates from NGI database which is characterized by a good temporal quality. While the second objective aims to extract semantic information about the building function from the cadastral database which represents the highest semantic resolution.

As we have already stated, adoption of CityGML as a reference standard for 3D modeling is motivated by its semantic richness, its support of multi resolution features by means of different levels of details and also for its interoperability. Likewise, CityGML is recommended by the European Inspire directive, especially through the definition of a 3D profile for buildings compatible with CityGML. However, the complexity of CityGML makes it hard to implement all its specifications. For our experiment, we have adopted a basic version of a 3D building model based on the building thematic module of CityGML. Constrained by the nature of data sources which is limited to 2D geometry augmented by information about the z coordinate, we have adopted a 3D model where buildings are represented by flat roofs and distinguishable thematic surfaces (Ground/Wall/Roof), as described in CityGML. This can be considered as an intermediate solution between LOD1 and LOD2, so designed by LOD1.5. The 3D geometric model is completed by adding attributes to each building such as ID, function, usage and status. The adopted 3D conceptual model is presented in Fig 5.

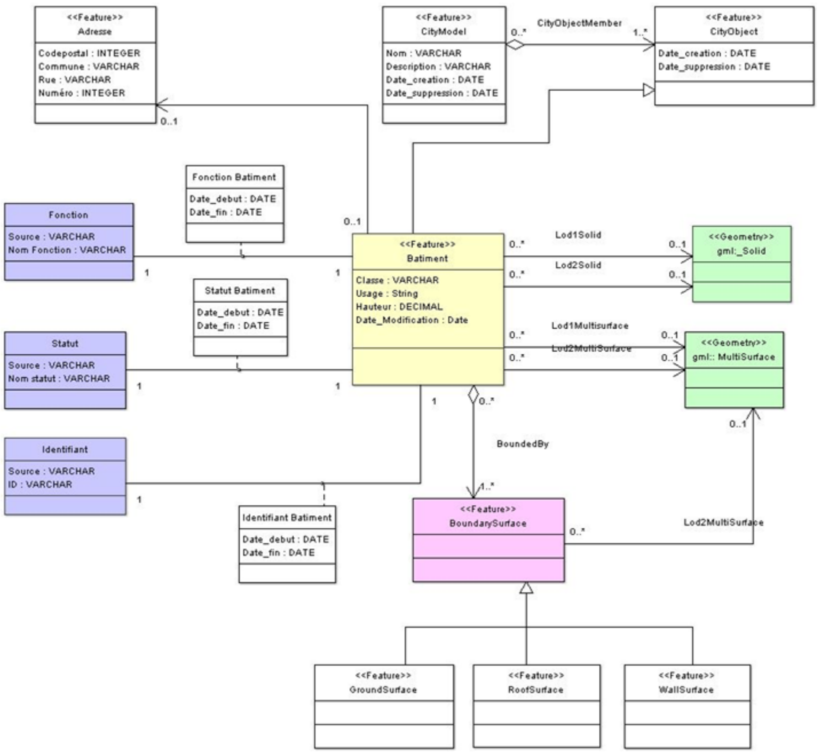

Fig. 5. The adopted 3D model

5.1 Databases Preintegration

The databases to be integrated are heterogeneous in several respects. Their combination highlights both geometric and semantic conflicts because of different specifications and various modeling approaches. The preintegration consists of identifying and reducing such dissimilarities through a good understanding of the content of each database and the mapping between models to show similarities and possible correspondences. As a preliminary task, it was necessary to unify coordinate systems, to reconstruct topology and to identify and classify the geometric and semantic conflicts.

5.2 Data matching

The process of data matching consists of the identification and declaration of correspondences between the elements of the schemas and the geometrical instances of the databases. Different approaches of data matching have been investigated in ((Yan-ling et al. 2016); (Walter et al.,1999); (Quddus et al. 2007), etc). There can be classified to three classes: Geometric, semantic ones based on semantic and those combining both geometric and semantic criteria.

To do the experiment, we have adopted an approach based on the surface distance as a geometric criterion for data matching to which the "building function" attribute has been added as a semantic criterion. The Wu-Palmer distance was adopted to calculate a similarity index between concepts to be aligned, which is added as an additional criterion with the geometric one to enhance the matching process (Fig. 6)

The surface distance is introduced by Vauglin (1997), which is defined as follows:
Let S (A) and S (B) be the respective measurements of the areas of the entities A and B, with S (A) ≠ 0 and S (B) ≠ 0. Note Δ the operator of the symmetric difference (A Δ B = A \ B + B \ A, with A\B is the complement of B in A). The surfacic distance is then given by: $Ds(A,B) = \frac{S(A\Delta B)}{S(A\cup B)}$

The Wu-Palmer distance establishes a value of the similarity between two concepts of the same taxonomy to the distance to their smallest generalizing common. Indeed, if C is the smallest generalizing common of two concepts C1 and C2, prof (C) the number of arcs that separate C to the root of the taxonomy, and $prof_C$ (Ci) the number of arcs that separate Ci from the root of the taxonomy, passing through C, then the measure of Wu-Palmer similarity between C1 and C2 is given by the following formula: $sim\ (C1, c2) = \frac{2*prof\ (C)}{profc\ (C1)+profc(C2)}$

In the following, we note DBRef: the reference database and DBComp: the comparison one. That is to say, for the instances of the DBRef we look for potential homologues in DBComp.

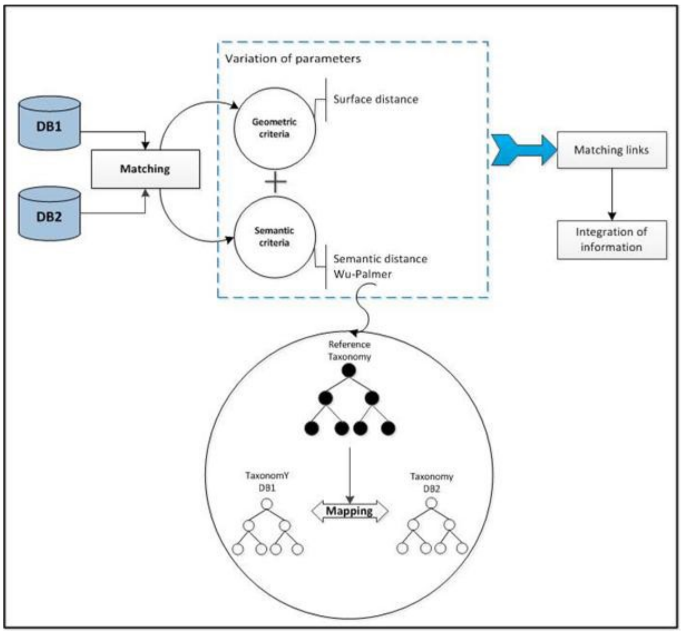

Fig. 6. The Matching process

We carried out several tests in order to evaluate the impact of geometric and semantic criteria on matching results, by varying the weight attributed to each parameter. As a preliminary test, a threshold of 5m was adopted to remove parasite intersections. Then, with a value of the inclusion function less than 10%, they are filtered in order to eliminate insufficient intersections.

In our case, the Walloon database is adopted as a DBRef. Two data matching tests were realized. The first one (PICC-NGI) uses the PICC database as a DBRef and the NGI database as a DBComp in order to extract the geometric updates. The second data matching (PICC-cadaster) uses the cadastral database as a DBComp in order to extract semantic information associated to the building function attribute. The PICC-cadaster matching combines geometric and semantic criteria. While the PICC-NGI matching is based only on geometric criteria.

The output of the data matching algorithm consists of matching links with different cardinalities ([1-1], [1-n], [n-1], [n-m]). The results are evaluated according a quantitative assessment of the detection quality by adopting a manual matching as a reference.

Based on the number of True positives (TP), False Positives (FP) and False Negatives (FN), different quality measures can be calculated. Table 1 presents some commonly used with data matching, such as detection rate, omission rate, over-detection rate, branch factor, miss factor and quality rate. These quality measures are calculated for each matching test with respect to the reference manual matching (Table 2 and Table 3).

Table 1. Quality measures for data matching (Frédérique, 2008)

Quality measure	Value
Detection rate	94%
Omission rate	5%
Over-detection rate	7%
Branch factor	0.08
Miss factor	0.06
Quality rate	87%

Table 2. Data matching quality (PICC-Cadaster)

Quality parameter	Equation
Detection rate	100.(TP/TP+FN)
Omission rate	100.(FN/TP+FN)
Over-detection rate	100.(FP/TP+FP)
Branch factor	FP/TP
Miss factor	FN/TP
Quality rate	100.(TP/TP+FN+FP)

Table 3. Data matching quality (PICC-NGI)

Quality measure	Value
Detection rate	96%
Omission rate	3%
Over-detection rate	1%
Branch factor	0.01
Miss factor	0.03
Quality rate	95%

In the case of matching PICC-cadaster, we note a quality rate of 87%. 7% of resulting links are over-detected. These links have the cardinality of [1-n] and are generated where several candidates have the same values or close values for both the surfacic and the semantic distance. These links can be reduced by increasing the value of the inclusion function.

In the case of matching PICC-NGI, the quality rate reaches 95%, with only 1% of over detection rate. This can be explained by the difference of resolutions between the two databases that had a good impact on the detection quality.

The resulting links are inspected and filtered before to be validated in order to maintain only those with a value of 0.4 of the surfacic distance (60% of inclusion) in order to eliminate some links eventually generated between objects semantically similar but with insufficient geometric proximity.

We have then used the resulting links of data matching for information fusion. In this step, we have considered the links resulting from the filtering process as "true" links to analyze their cardinality for an information fusion. While invalid or "uncertain" links are to be validated by an external source of information.

We have conducted two types of analysis in accordance with the predefined objectives of the data matching, which consist of extracting semantic information (from cadastral database) and geometric updates (from NGI database). The strategy of information fusion was based on the analysis of the cardinalities of matching links. Semantic information is transferred on the geometric position of the reference object according to the links cardinality, resulting from PICC-cadaster matching. On the other hand, the analysis of the results of PICC-NGI matching has been of a particular interest in the detection of geometric updates which have been exported to the PICC database.

Finally, we have to say that the information fusion is not a trivial task. In case of multiple links (with [n-m] cardinality), an additional source of information is required in order to validate some hypothesis.

5.3 Data integration and reengineering

In accordance with the predefined objectives of the data matching, the strategy of information fusion was based on the analysis of the cardinalities of matching links. Semantic information is transferred on the geometric position of the reference object according to the links cardinality, resulting from PICC-cadaster matching. On the other hand, geometric updates have been transferred to the reference database.

Data reengineering consists of defining explicit rules to restructure the initial schemas and to translate data according the specifications of the target system. Indeed, the layer resulting from data fusion is to be converted from a 2D geometry to a 3D geometry compatible with CityGML schema.

As a first step, we had to interpolate the buildings' heights for each data set. To accomplish this task, a Digital Surface Model (DSM) and a Digital Terrain Model (DTM) were established. The DSM or exactly a "Digital Cornice Model" (DCM) was generated from the contours of building's roofs. After rasterization of the two TINs surfaces and the building contours roofs, statistical computing based on the DTM and the DSM was used to extract the height of each building (Fig. 7)

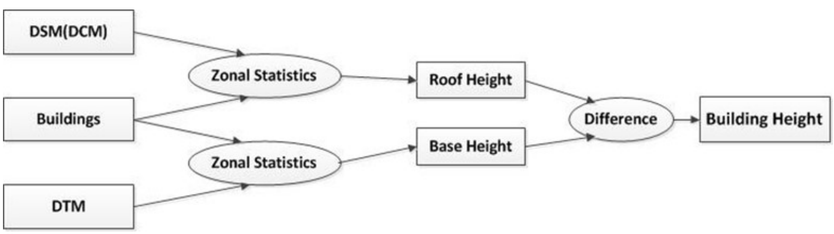

Fig. 7. Interpolation of building heights

We have adopted extrusion as a simplest way to construct 3D buildings while topological relationships between the footprints were not considered. Building blocks are obtained from a 2D GIS layer. The value of their height is assigned from the previous process of interpolation. Buildings' geometries are then 3D-reconstructed in 3DSMax program. The result is a 3D geometry structured in different layers associated to thematic surfaces of buildings as defined in the building module of CityGML (Roof Surface; Ground Surface and Wall Surface).

After the 3D reconstruction of buildings, FME program is used to perform the conversion to the CityGML schema. FME is a spatial ETL (Extract, Transform and Load) application concept which provides unlimited flexibility in data model transformation, translation and integration by dealing with a large number of formats. The result is a CityGML file which can be imported in an Oracle database instance implemented using "3DCityDatabase" for Oracle, a free and open source 3D geo database to store, represent, and manage virtual 3D city models, which schema is compatible with CityGML. The resulting 3D model is presented in Fig. 8.

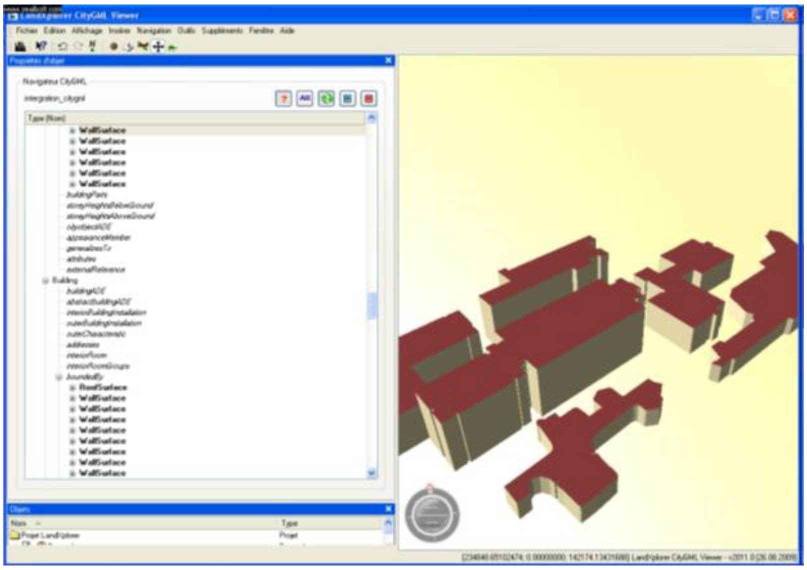

Fig. 8. The resulting 3D model of buildings with CityGML schema

6. Conclusion & discussion

In this paper, we have presented a process of data integration in order to establish a 3D building database based on CityGML. CityGML has been seen as a reference standard for the 3D model in order to ensure interoperability between 3D city models.

Our work can be considered as a building block to integrate existing 2D/2.5D data into a 3D database which schema is compatible with CityGML in order to establish a basic 3D building model. Investigations are still needed to deal with the enrichment of these 3D models by considering other data sources and defining specifications for future 3D data production.

Based on our investigations, we can state that a promising issue about 3D modeling is the collaborative work. In fact, many actors should collaborate effectively to build a collaborative 3D model to respond to emergent challenges related to 3D geospatial data. In order to avoid duplicated efforts and expenses, data provided by many contributors must be integrated to develop reference 3D data sets. Doing so, data interoperability can be greatly improved. To reach the objective of establishing a collaborative 3D reference model, a 3D geographic data framework must be established through agreement on content and specifications (Hajji and Billen, 2012). The framework will constitute a collaborative "datum" on which different actors can build by adding their own details and compiling other data sets. In short, building a 3D collaborative model would be a promising solution firstly, to anticipate interoperability and data consistency problems at a technical perspective and secondly to establish common specifications for 3D data co-production (Hajji, 2013).

References

- Akmalia, R., Setan, H., Majid, Z., Suwardhi, D., & Chong, A. (2014). TLS for generating multi-LOD of 3D building model. In IOP conference series: Earth and environmental science (Vol. 18, No. 1, p. 012064). IOP Publishing.

- Atila, U., Karas, I. R., Rahman, A. A., (2013). A 3D-GIS Implementation for Realizing 3D Network Analysis and Routing Simulation for Evacuation Purpose. In *Progress and New Trends in 3D Geoinformation Sciences* (pp. 249-260). Springer Berlin Heidelberg.

- Aydinoglu, A. C., Quak, W.,Yomralioglu, T., (2009). Some Spatial Data Management Issues towards Building SDI. In *International Workshop on Spatial Information for Sustainable Management of Urban Areas, FIG Commission 3 Workshop*, Mainz, Germany (pp. 2-4).

- Benner J., Geiger A., Häfele K.H., 2010. Concept for building licensing based on standardized 3D Geo information. In Proc. of the 5th International 3D GeoInfo Conference.

- Biljecki, F.,Ledoux, H.,Stoter,J. (2016). An improved LOD specification for 3D building models. Computers,Environment,and Urban Systems,vol. 59,pp. 25-37.

- Biljecki, F., Stoter, J., Ledoux, H., Zlatanova, S., & Çöltekin, A. (2015). Applications of 3D city models: State of the art review. ISPRS International Journal of Geo-Information, 4(4), 2842-2889.

- Breunig and Zlatanova, S. (2011). 3D Geodatabase research: Retrospective and future directions. Computers & Geosciences

- Deng, Y., & Cheng, J. C. (2015). Automatic transformation of different levels of detail in 3D GIS city models in CityGML. *International Journal of 3-D Information Modeling (IJ3DIM)*, *4*(3), 1-21.

- Emgard, K. L., Zlatanova, S., (2007). Design of an integrated 3D information model. Urban and regional data management: UDMS annual, 143-156.
- Frédérique, B. (2008). *Saisie photogrammétrique multireprésentation de bâtiments: Une approche Semi-Automatique Initialisée et Supportée par l'Intervention humainE*. PHD Thesis, Faculty of Graduate Studies, Laval University.
- G. Gröger, T. Kolbe, et al., (2008). OpenGIS City Geography Markup Language (CityGML) Encoding Standard (OGC 08-007r1)
- Goodchild, M. F., (2010). Twenty years of progress: GIScience in 2010. *Journal of Spatial Information Science*, (1), 3-20.
- Gröger, G., Plümer, L., (2012). CityGML–Interoperable semantic 3D city models. *ISPRS Journal of Photogrammetry and Remote Sensing*, 71, 12-33.
- Hajji, R. (2013). *Vers une méthode de conception de SIG 3D collaboratif*. PhD. Liege University.Belgium.
- Hajji, R., and Billen, R., 2012. Towards a collaborative and interoperable 3D Building database–A case study in Walloon region. In Usage, Usability, and Utility of 3D City Models–European COST Action TU0801 (p. 03004). edp Sciences.
- Ibraheem A. T., Hassan H. A., Al-Husain M. H. A. (2012). Integrating ACAD with GIS for Civil Engineering Applications. *Journal of Software Engineering and Applications*, 5(3), 138-146.
- Isikdag U., Zlatanova S., 2009. A SWOT analysis on the implementation of Building Information Models within the Geospatial Environment. Urban and Regional data Management, *UDMS Annuals*, 15-30. El-Mekawy, M., Östman, A., 2010. Semantic Mapping: an Ontology Engineering Method for Integrating Building Models in IFC and CITYGML. Proceedings of the 3rd ISDE Digital Earth Summit, 12-14.
- Karimi H. A., Akinci B., 2010. *CAD and GIS Integration*, Taylor & Francis.2010
- Lowner, M.-O., Benner, J., Groger, G., & Hafele, K.-H. (2013). New Concepts for Structuring 3D City Models
- Nagel, C., Stadler, A., Kolbe, T.H., 2009. Conceptual requirements for the automatic reconstruction of building information models from uninterpreted 3D models. In Academic Track of Geoweb 2009 Conference, Vancouver.
- Quddus, M. A., Ochieng, W. Y., & Noland, R. B. (2007). Current map-matching algorithms for transport applications: State-of-the art and future research directions. Transportation research part c: Emerging technologies, 15(5), 312-328.
- Rahman, A.A, Pilouk, M. (2008). *Spatial Data Modeling for 3D GIS*. ISBN 978-3-540-74166-4 Springer Berlin Heidelberg New York.
- Robles-Ortega, M. D., Ortega, L., Feito, F.R., (2012). Design of Topologically Structured Geo-database for Interactive Navigation and Exploration in 3D Web-Based Urban Information Systems. *Journal of Environmental Informatics*, 19, (2), 79-92.
- Sheeren, D., Mustière, S., Zucker, J-D., (2004). How to Integrate Heterogeneous Spatial Databases in a Consistent Way? *Conference on Advanced Databases and Information Systems*, Budapest, September 2004, p. 364-378.
- Stoter, J., van den Brink, L., Vosselman, G., Goos, J., Zlatanova, S., Verbree, E., ... & Thorn, S., 2011. A generic approach for 3D SDI in the Netherlands. In *Proceedings of the Joint ISPRS Workshop on 3D City Modeling & Applications and the 6th 3D GeoInfo Conference* (pp. 26-28).

- Van den Brink, L., Stoter, J., Zlatanova, S., (2013). Establishing a national standard for 3D topographic data compliant to CityGML. *International Journal of Geographical Information Science*, 27(1), 92-113.

- Vauglin F. (1997). Modèles statistiques des imprécisions géométriques des objets géographiques linéaires. Thèse de doctorat, Science de l'information géographique, Université de Marne-La-Vallée, 1997.

- Walter, V., & Fritsch, D. (1999). Matching spatial data sets: a statistical approach. International Journal of geographical information science, 13(5), 445-473.

- Yan-ling, H. A. O., Wen-jing, T. A. N. G., Yu-xin, Z. H. A. O., & Ning, L. I. (2016). Areal feature matching algorithm based on spatial similarity. Acta Geodaetica et Cartographica Sinica, 37(4), 501-506.

- Zlatanova, S. Stoter, J., Isikdag U., (2012). Standards for Exchange and Storage of 3D Information: Challenges and Opportunities for Emergency Response. *Proceedings of the 4th International Conference on Cartography & GIS*, Volume 2, Albena, June 2012, pp. 17-28