Modeling Concepts for Consistency Analysis of Multiple Representations and Heterogeneous 3D Geodata

Susanne Becker, Volker Walter, and Dieter Fritsch

Abstract In a dynamic world where the steadily increasing demand for up-to-date geodata drives the continuous acquisition of three-dimensional (3D) data, appropriate systems for managing and analyzing the resulting data become more and more important. Efficient solutions for handling multiple representations and data heterogeneity are of special significance. Existing geoinformation systems are still not able to cope with the huge diversity of geodata. Available approaches and systems that apply merging processes in order to generate one single representation for each real-world object are not practicable any more. Thus, our goal is a hybrid 3D geoinformation system that allows for integrated management of heterogeneous and multiply-represented geodata. Our concept is hybrid with respect to data given in different data models, dimensions, and quality levels. Multiple representations and data inconsistency can be handled through the explicit modeling of geometric correspondences.

Keywords GIS · Integration · Modeling · Three-dimensional · Consistency

1 Introduction

Dealing with multiple representations and heterogeneous data are important issues on the way toward full interoperability in geoinformation systems. Multiple representations result from the fast increasing availability of geodata. On the one hand, this flood of geoinformation implies immense potential for solving various problems. On the other hand, due to the multitude of different sensors, algorithms, and modeling concepts used for data acquisition and processing, such geodata is highly complex and heterogeneous—posing a big challenge when the data has to

S. Becker (✉) · V. Walter · D. Fritsch
Institute for Photogrammetry, University of Stuttgart, 70174 Stuttgart, Germany
e-mail: susanne.becker@ifp.uni-stuttgart.de

V. Popovich et al. (eds.), *Information Fusion and Geographic Information Systems (IF&GIS 2013)*, Lecture Notes in Geoinformation and Cartography,
DOI: 10.1007/978-3-642-31833-7_6,

be evaluated together. Data heterogeneity generally goes back to structural aspects concerning conceptual data modeling, as well as geometric and topological issues.

In terms of structural heterogeneity in conceptual modeling, Bishr (1998) distinguishes between semantic, schema, and syntactic heterogeneity (Gröger and Kolbe 2003). Semantic heterogeneity arises when dissimilar ways of understanding real-world phenomena lead to different object abstractions. Schema heterogeneity, however, denotes structural differences in the modeling concept. For instance, the same object property could be modeled as a class in concept A but as an attribute in concept B. Syntactic heterogeneity is related to different geometric data models: The two-dimensional (2D) world is mainly based on raster and vector representations; typical data models for 2.5D surfaces are grids or TINs; three-dimensional (3D) solids can be described by voxels, boundary representations (BRep), mathematical definitions such as parametric instancing, half-space modeling, or constructive solid geometry (CSG) and cell decomposition.

Geometric and topological heterogeneity is related to various aspects, particularly different reference systems and data quality. Due to the applied sensors and their configuration during measurement, data sets can differ significantly in accuracy, resolution, density, and completeness. Inconsistencies may occur when the integration of various data sets lead to spatial intersections or interpenetrations of different geoobjects, or when there remain gaps between geometries that in fact are adjacent.

Considering all these aspects, merging multiple representations in order to achieve a consistent view on the data is obviously impracticable. Solutions considering today's challenges should aim at an integrated management of multiple representations. This, however, requires geoinformation systems that allow for the explicit modeling of multiply represented data given in various geometric data models, dimensions, and quality levels. Moreover, an integrated management of multiple representations inevitably leads to inconsistencies, which have to be handled in an appropriate way. In this respect, traditional geoinformation systems that try to ensure full data consistency are not powerful and flexible enough.

In order to overcome the lack of appropriate concepts and methods for handling heterogeneous and multiply represented geodata, we propose an all-encompassing modeling concept that extends an existing International Organization for Standardization (ISO) standard in such a way that it is hybrid in the sense of data model, dimension, and quality. Our data model is designed to be an appropriate basis for a powerful and flexible 3D geoinformation system: powerful because it provides the basis for efficient consistency analyses and updating processes, and flexible because it supports multiple representations and is able to cope with structural as well as geometric and topological data heterogeneity. Visualization aspects and the modeling of semantics are not taken into account here.

The chapter is organized as follows: After an overview of related work in Sect. 2, our hybrid data model will be presented in Sect. 3. The explicit modeling of multiple representations will be introduced in Sect. 4. Section 5 will demonstrate how our modeling concepts can be used for efficient consistency analyses. Finally, Sect. 6 will conclude the paper.

2 Related Work

Data heterogeneity is a complex and multifaceted topic. Thus, approaches dealing with heterogeneous geodata usually focus on certain subproblems. They try to overcome either structural heterogeneity covering semantic, schema, or syntactic issues or geometric and topological heterogeneity. An overview of approaches that address structural heterogeneity will be given in Sect. 2.1. Section 2.2 will present current work focusing on data inconsistency caused by geometric and topological heterogeneity. Resulting consequences for our work will be discussed in Sect. 2.3.

2.1 Data Heterogeneity

Concerning structural aspects, the first investigations on hybrid data models and analysis methods go back to the 1980s; however, they only cover the 2D world (Fritsch 1988). An integrated view of hybrid 3D data has only been a topic of research for a few years. A step in this direction was taken by Dakowicz and Gold (2010), who went beyond pure 2D representations by suggesting a unified spatial model for 2D and 2.5D data. Existing approaches that are also able to handle 3D data are generally tailored to specific applications; thus, they just address subproblems, such as the combination of 2D and 3D building data (Inhye et al. 2007), the merging of TINs and grids for the representation of digital elevation models (Proctor and Gerber 2004), or the handling of CSG- and BRep-models in computer-aided design (Stekolschik 2007). Lee and Zlatanova (2008) propose a 3D data model especially suited for emergency response. Here, neighborhood relations are explicitly modeled through graph models, allowing for efficient routing algorithms; the geometric part of the data model is limited to BRep-representations, though. The same restriction holds for the slice representation introduced by Chen and Schneider (2009) as a general data representation method for 3D spatial data.

The approaches mentioned so far address structural data heterogeneity with respect to very specific application scenarios and, thus, are not suitable for general use. An application-independent conceptual framework to integrate discrete objects and continuous field-based objects on a logical level is given by Voudouris (2010); however, the study does not consider how different implementations of such object- and field-based models (e.g., vector and raster) can be managed to overcome the inherent syntactic data heterogeneity.

In principle, standards are indispensable when interoperability problems have to be avoided. The Open Geospatial Consortium (OGC) is one of the driving forces for the development of standards. OpenGIS is the brand name for standardization processes under the umbrella of OGC. A lot of OpenGIS specifications have already become an ISO standard, such as the OGC Topic 1 “Feature geometry,” whose specifications and concepts can also be found in the ISO 19107 Spatial

Scheme standard. Focusing on the description of vector data only, ISO 19107 comprises geometric and topological modeling concepts for 3D objects (Herring 2001). Because only BRep is supported, syntactic data heterogeneity remains a problem (Gröger and Kolbe 2003). ISO 19107 is based on set, theory. Each geometry is interpreted as a topological point set, which is the basis for specifying spatial comparison operators.

The Special Interest Group 3D (SIG 3D) proposed a specification for city models, CityGML, which is based on ISO 19107 (Kolbe et al. 2005). CityGML means a considerable advance to the interoperability of 3D city models. Nevertheless, a lossless integration of data that follows the CSG modeling approach still is not possible because CSG concepts are not supported. It is also not possible to integrate parametric instancing as it is often used for modeling frequently occurring similar objects.

Seen from a conceptual point of view, data integration is feasible without interoperability problems when object representations follow the same modeling standard. However, this only holds true if, additionally, the data are consistent in terms of geometric and topological aspects. Generally, multiple representations cannot be expected to be consistent as to accuracy, completeness, level of detail, etc. Thus, standardizations are only of limited use for multiply represented geodata.

2.2 Consistency and Multiple Representations

Geometric and topological data heterogeneity inevitably leads to inconsistencies in merged data sets. Gröger and Plümer (2011) support consistency analyses in 3D city models by specifying axioms for topological components and their aggregations. For this purpose, the city model is topologically interpreted as a complete and unique 3D tessellation where each geometric object is represented exactly once. Multiple representations are not supported. However, detecting and managing multiply represented objects plays an important role in geoinformation systems, especially when different data sets are to be combined. In the 2D world, a number of approaches have been developed; each of these approaches focuses on specific data types. For instance, Walter (1997) proposed a method for the matching of street data from different sources. Based on this, Volz and Walter (2004) realized the integration of multiply represented 2D vector data on the schema level. Although the range of approaches for identifying and processing multiply represented 2D geodata is wide, the situation is different for 3D data. The first ideas for analyzing the consistency of selected 3D geometries have been presented in recent years. For example, Peter (2009) compares geometric properties of planar 3D faces to estimate the consistency of different building representations. However, there is still a considerable need for research in this area.

2.3 Consequences for Our Work

The review of existing approaches dealing with heterogeneous geodata reveals a number of yet unsolved challenges on the way toward full data interoperability. The vast majority of approaches present application-specific solutions for a rather narrow range of different data types, data models, or quality levels; an overall modeling concept for arbitrary geodata is still missing. Additional problems and limitations result from the separate consideration of structural heterogeneity on the one side and geometric and topological heterogeneity on the other side. For example, although CityGML aims at interoperability on a semantic and syntactic level, the explicit treatment of geometric and topological heterogeneity is neglected; consistency analyses are not supported. However, a full interoperability and, no less important, a sustainable management of geodata, which is a basic requirement for efficient analyses and updating processes, necessarily demands for an integrated view on all heterogeneity aspects—structural as well as geometric and topological ones. Present and future challenges in the field of geoinformation demonstrate the urgent need for systems that are able to manage heterogeneous and multiply-represented geodata and, furthermore, support consistency analyses.

3 Hybrid Data Model

We introduce an application-independent modeling concept that is hybrid in the sense of structural and geometric plus topological aspects. Our data model basically builds on two modeling decisions. The first one, which will be described in Sect. 3.1, is related to the idea of expressing arbitrary object representations through geometric elements that are part of all existing data models. The second modeling decision refers to the aim of creating as much interoperability as possible and, thus, results in applying the standard ISO 19107. The geometric primitives specified in ISO 19107 are particularly appropriate to model discrete objects. However, they can also be used to define the geometric basis (e.g., sampling points) for continuous, field-based objects (Andrae 2009) and, thus, pave the path to a unified, concept for modeling both discrete object-based phenomena and continuous field-based phenomena. The basic modeling principles of ISO 19107 as well as our extensions to the standard will be explained in Sect. 3.2.

3.1 Hybrid Core

In order to overcome syntactical heterogeneity, we base our data model on fundamental modeling elements that are part of the most relevant existing geometric data models. We introduce the term *hybrid core* to denote such common modeling elements. It appears that a *general* hybrid core that is valid for all data models does not exist. To prove this, it is sufficient to compare the 2D vector format with the 2D raster representation. 2D vector data is modeled through points, lines, and faces. Because lines and faces, in turn, are described by sequences of points, the point turns out to be the basic modeling element of 2D vector data. The existence of a hybrid core for vector and raster data would implicate the point to be a basic modeling element of raster data, too. However, points cannot be expressed in the raster format in purely geometric terms. The explicit semantic modeling as point object is additionally required because—as a consequence from the approximating character of raster data—a raster cell could also represent a short line or a small surface.

Because a *general* hybrid core is not available, we create an *artificial* one based on the working hypothesis, which states that all modeling types considered so far (e.g., vector, raster, TIN, grid, voxel, cell decomposition, CSG, etc.) can be transferred to BRep. By internally creating BRep for all data sets, even syntactically inhomogeneous geodata can be reduced to a hybrid core comprising points, lines, surfaces, and solids. In the case of raster and voxel data, where each 2D or 3D cell is then described by its bounding lines or surfaces, respectively, this modeling concept is of course not efficient. However, according to fast advances in the development of high-speed processors and parallel computing, it seems reasonable to ignore performance issues for now. Efficient access structures can be added to the model at a later stage.

3.2 Extension of the Standard ISO 19107

In order to ensure as much interoperability as possible, we build our modeling concept on the ISO 19107 standard. ISO 19107 is a widely accepted standard for the modeling of geometric and topological aspects of a real-world phenomenon (Andrae 2009). Based on BRep, it is appropriate to describe 2D and 3D vector data as well as TINs and grids. We propose several standard compatible extensions that open the standard to further geometric representations. Here, the focus is on approximating data models such as raster and voxel, and on the CSG modeling approach.

Figure 1 shows our data model in UML notation; explanations will be given in the following sections: Essential modeling principles of ISO 19107, the basis of our data model, will be described in Sect. 3.2.1 (Fig. 1 presents corresponding object classes in light gray). The standard compatible extensions for raster and voxel data will be given in Sect. 3.2.2 (highlighted in orange, red frames (horizontally hatched)), while Sect. 3.2.3 will show how the CSG concept (highlighted in orange, blue bold frames (horizontally hatched)) can be integrated in ISO 19107. By means of the object classes and corresponding associations colored in green (diagonally hatched), Fig. 1 illustrates how various geometric data representations can be expressed by our data model.

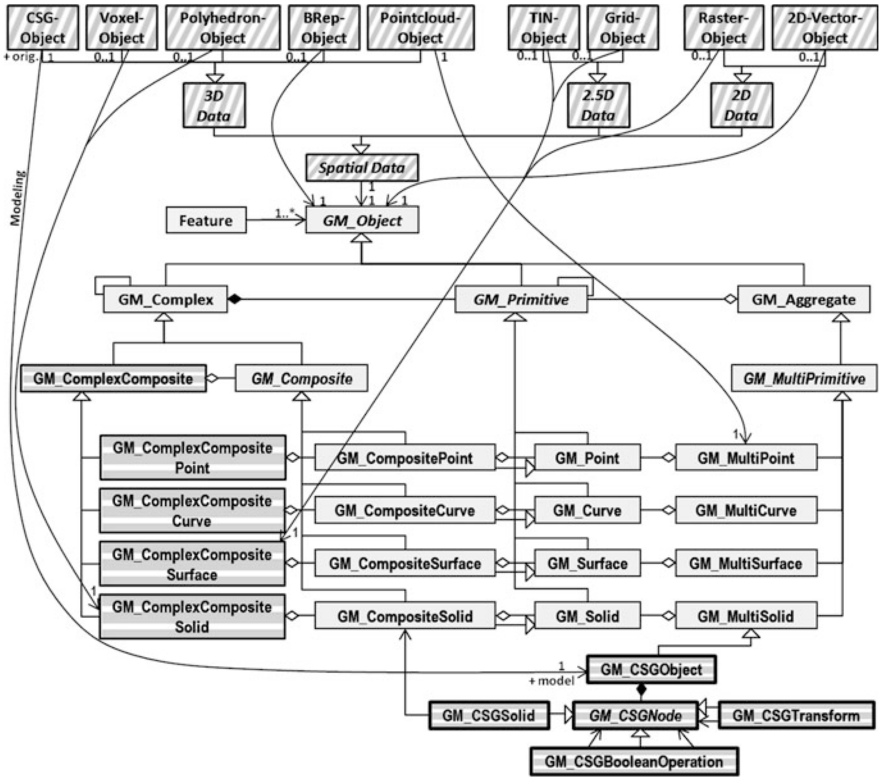

Fig. 1 Hybrid data model

3.2.1 Basic Modeling Principles

ISO 19107 defines GM_Object as a base class for the geometric properties of all geoobjects. An instance of GM_Object is either a GM_Primitive, a GM_Aggregate, or a GM_Complex. Specializations of GM_Primitive are the classes GM_Point, GM_Curve, GM_Surface, and GM_Solid. These geometric primitives cannot be divided into further primitives and, thus, represent basic elements. Instances of the class GM_Aggregate are unstructured collections of geometries free of any topological restrictions. Aggregates whose components all belong to the same primitive type are elements of the class GM_MultiPrimitive.

In contrast to GM_Aggregate, GM_Complex offers an opportunity to combine geometric elements in a structured way. Topological constraints ensure these elements to be disjoint and not self-intersecting; they are allowed to touch each other, though. A complex belongs to the class GM_Composite if the following additional conditions are fulfilled: (1) all components of the complex are of the same primitive type; (2) the complex is isomorphic to a primitive. Important specializations of GM_Composite are GM_CompositeCurve, GM_CompositeSurface, and GM_CompositeSolid.

As mentioned above, ISO 19107 additionally allows for the explicit modeling of a geoobject's topological properties by separate classes. To simplify matters for now, we do without an explicit topological modeling. Topological properties can be derived anyway when geoobjects are modeled as instances of the class GM_Complex. A geometric complex describes topology implicitly because ISO 19107 defines that—in contrast to primitives and aggregates that represent open sets—a complex contains its components plus the boundary of each component.

3.2.2 Extensions for Raster and Voxel Data

According to our working hypothesis, a raster representation of a geoobject can be interpreted as a composition of single surface elements, in which each surface element corresponds to one pixel and is described by its bounding lines. Figure 2 shows an exemplary 2D geoobject in both raster (Fig. 2a) and boundary representation (Fig. 2b). In order to emphasize the different characteristics of these two concepts, pixels are illustrated in black, surface elements in gray with black boundaries. Due to the properties and topological relations of raster cells (not self-intersecting, disjoint), such a composition of surface elements meets the requirements of a GM_Complex. But, modeling a raster object as a general complex means losing knowledge about important geometric properties because a complex does not know about its components' primitive types: ISO 19107 does not specify or restrict which primitive types may occur in a complex; even a mixture of dissimilar types is allowed.

Modeling a raster object instead as an instance of GM_CompositeSurface, which is a specialization of GM_Composite and, thus, also of GM_Complex, would preserve the knowledge about occurring primitive types. However, as will be shown by the examples in Fig. 2c, GM_CompositeSurface cannot express raster objects of arbitrary shape. The reason is that a composite is defined to be isomorphic to a primitive; consequently, a composite surface—here, the union of various raster cells—has to be isomorphic to a single-surface primitive. As ISO 19107 requires a surface primitive to be simple, i.e., free of self-intersections and self-touches; only those raster objects can be modeled as a valid composite surface whose raster cells each have at least one edge in common with another raster cell. Although this is true for Fig. 2c1, raster objects similar to the example in Fig. 2c2 cannot be modeled as composite because the outer boundary of the merged cells touches itself.

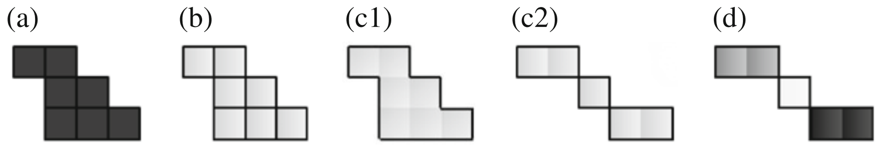

Fig. 2 **a** Raster object, **b** Raster object interpreted as BRep, **c1** and **c2** Raster objects modeled as GM_CompositeSurface, **d** Raster object modeled as GM_ComplexCompositeSurface

To overcome this problem, the new class GM_ComplexComposite is introduced as a specialization of GM_Complex. An instance of this class is a complex of several composites, which can be of different composite types. Restrictions forcing these composites to be of identical primitive type are realized through the specializations GM_ComplexCompositePoint, GM_ComplexCompositeCurve, GM_ComplexCompositeSurface, and GM_ComplexCompositeSolid. Based on this extension to the data model, it is now possible to model raster objects of arbitrary shape. The class appropriate for this purpose, GM_ComplexCompositeSurface, even allows for the modeling of completely unconnected raster cells or raster configurations in which cells are connected through just a corner, as is the case in Fig. 2c2. As illustrated in Fig. 2d, parts of the object that are isomorphic to a single surface are modeled as instances of GM_CompositeSurface; all together, they can then be interpreted as a complex of three composite surfaces, i.e., as an instance of the class GM_ComplexCompositeSurface. Extensions for the modeling of voxel representations follow analogous considerations. The new object class introduced for this purpose is called GM_ComplexCompositeSolid.

3.2.3 Extensions for CSG Data

In principle, CSG data can be converted into BRep by determining the visible bounding faces. Doing so, however, implies the loss of information on the construction process and geometric conditions of the CSG object (Gröger et al. 2005). Such information can be relevant for updating purposes.

We integrate the CSG concept in the data model through the new object class GM_CSGObject. Derived from the aggregate GM_MultiSolid, this class allows its components to overlap and penetrate each other, which is a characteristic property of CSG objects. By means of the so-called CSG node, realized through the class GM_CSGNode, the hierarchical structure of the CSG construction process can be modeled. GM_CSGNode serves as a base class to define transformations, Boolean operations, and CSG solids, the constructive elements of a CSG object. A Boolean operation, for example, refers to two nodes to which it is applied. Transformations are modeled accordingly. A CSG solid refers to an instance of GM_CompositeSolid, which ensures that the solid's boundary is a part of the object.

Our object-oriented way of modeling CSG objects makes it possible to completely hide their constructive design from the rest of the standard. Special analysis methods for CSG objects can be introduced without changing the standard.

4 Management of Multiple Representations

The hybrid data model proposed in Sect. 3 can cope with structural heterogeneity; data of different dimensions and geometric representations can be handled, analyzed, and visualized together. In order to further increase the flexibility of our data model, we introduce concepts for the explicit modeling of multiply represented data in consideration of geometric and topological data heterogeneity (Sects. 4.1 and 4.2).

4.1 Modeling Concept for Multiple Representations

An arbitrary geoobject, which is called a *feature* in our data model, can be realized through one or more representations, each of them modeled as an instance of GM_Object. These instances actually do not need to cover the geoobject completely, but instead can also describe only parts of the object. Thus, on the one hand, our modeling concept provides the possibility to manage multiply represented geoobjects. On the other hand, it is also feasible to combine various object parts to one geoobject, even if these object parts stem from very different geometric representations (e.g., from a TIN mesh and a voxel representation).

However, an efficient usage, analysis, and interpretation of the data is only possible if geometric equivalences between different object representations are known, i.e., if it is known which geometry of one representation corresponds to which geometry of another representation of the same geoobject. In the following, we will denote such geometric correspondences between different object representations as *hybrid identities*.

Assuming an ideal world, in which coordinates of corresponding object representations coincide exactly, hybrid identities are given implicitly through incident geometries. As an example, Fig. 3 (left) depicts several representations of a simple building: a 3D vector representation of the building's solid, the 2D vector outline, a raster representation of the building's footprint, and a 3D point cloud observed at one building face. Because the boundaries of these representations exactly match with each other, corresponding geometries can automatically be derived by means of geometric comparisons.

Such an ideal situation illustrated in Fig. 3 (left) is a special case that can only occur as result of specific conversions or when one representation has been created based on another (e.g., a 3D solid through extruding a 2D outline). In practice, we usually face geodata that are geometrically and topologically heterogeneous due to

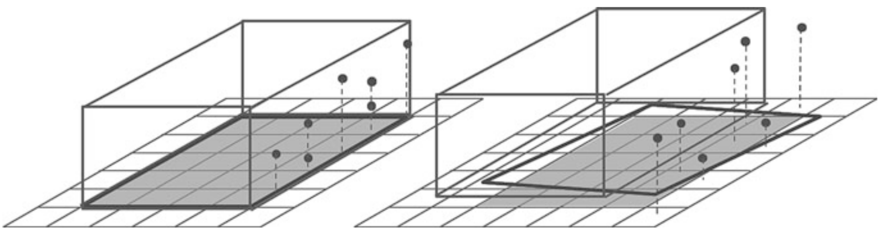

Fig. 3 Multiple representations of a building in an ideal, consistent, and error-free world (*left*) and in the real-world (*right*)

inaccuracies, generalization processes, or incomplete data acquisition. As a consequence, multiple object representations derived thereof show significant discrepancies between corresponding geometries (Fig. 3 (right)). Thus, knowledge about hybrid identities is not given implicitly any more but has to be added explicitly instead. Details on modeling aspects and the possible usage of hybrid identities are described in Sect. 4.2.

4.2 Modeling Concept for Hybrid Identities

Figure 4a shows the concept we developed for the explicit modeling of hybrid identities. The concept goes beyond the modeling of purely geometric aspects because knowledge about correspondences and relations between different object representations is introduced. The class HybridIdentity is used for managing hybrid identities. Each hybrid identity refers to at least two mutually corresponding structures modeled as instances of the class HybridElement. Depending on whether such a *hybrid element* stands for a single primitive or is a collection of several primitives, it can be a *hybrid primitive*, a *hybrid complex,* or a *hybrid aggregate*. The way in which several hybrid primitives are combined to a hybrid complex or aggregate follows the basic modeling principles as proposed in Sect. 3.2.1. In order to avoid redundancy, a hybrid primitive does not contain an explicit geometric description but refers to an existing instance of the class GM_Primitive. Conversely, an instance of GM_Object refers to all hybrid identities in which it is involved.

The data model for hybrid identities is designed to offer as much flexibility as possible. Being modeled independently of each other, hybrid identities can be defined for either a whole object or components of it. Additionally, one and the same object or object part can belong to several hybrid identities. Based on the example of a multiply represented 2D line object, Fig. 4b–e demonstrates a small selection of the many possibilities to define hybrid identities. Figure 4b shows the

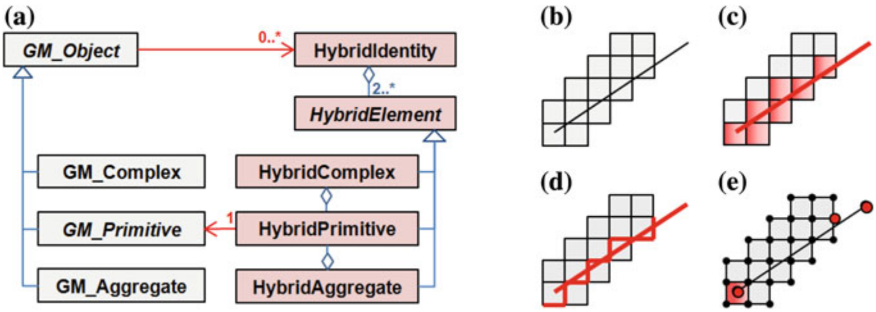

Fig. 4 **a** Data model for hybrid identities, **b** Raster and vector representation of a line object, **c**–**e** Exemplary definitions of geometric correspondences (*red*, *bold*)

two representations available for the 2D line object. The linear one (rep_A) stems from a 2D vector representation and is modeled as an instance of GM_Complex (here, consisting of a single line and its boundary). The areal one (rep_B) originates from raster data and is given as an instance of GM_ComplexCompositeSurface. Possible hybrid identities can be defined for the following geometries: the line of rep_A and a subset of the surface patches of rep_B (Fig. 4c); the line of rep_A and a sequence of lines bounding the surface patches of rep_B (Fig. 4d); the points bounding the line of rep_A and a single surface patch, as well as a boundary point of rep_B (Fig. 4e).

5 Hybrid Consistency

The concepts proposed in Sect. 4 provide the basis for efficient consistency analyses between arbitrary data sets. Through the combination of various GM_Objects to one feature (Sect. 4.1) and the definition of hybrid identities (Sect. 4.2), the pure geometric modeling is enriched by information on semantic entities, (i.e., knowledge about the relations between different object representations is introduced). This explicitly modeled knowledge about multiple representations and geometric correspondences provides the basis for consistency analyses. Integrated in our hybrid data model, consistency can now be evaluated and quantified even for highly heterogeneous object representations that stem from different data models and have different dimensions and quality levels. The traditional understanding of consistency as the lack of contradiction within a single data set or between two structurally homogeneous data sets consequently has to be extended to a so-called *hybrid consistency*. The definition of the term hybrid consistency is closely related to our hybrid data model and directly refers to the modeling concepts proposed for multiple representations and hybrid identities. Thus, it is possible to determine the degree of hybrid consistency between different object representations that describe the same real-world object either entirely or partially. Based on two exemplary scenarios, the following two sections will demonstrate how our modeling concepts can be used to evaluate hybrid consistency for entirely overlapping object representations on the one hand (Sect. 5.1), and partially overlapping or adjacent object representations on the other hand (Sect. 5.2).

5.1 Hybrid Consistency for Entirely Overlapping Object Representations

By "entirely overlapping object representations," we mean data sets that solely describe geometries of the same real-world object. Although these geometric descriptions do not have to cover the real-world object completely, they do not contain geometries from other neighboring real-world objects. Various aspects of our modeling concepts support the evaluation of hybrid consistency between multiple representations of one and the same real-world object.

First, the assignment of various GM_Objects to one feature object indicates all representations that are available for a specific real-world object. These multiple representations provide the geometric input for the consistency analysis.

Second, following the idea of expressing all data models through the geometric elements of a hybrid core, it is ensured that the object representations that are to be analyzed show structural homogeneity. According to our artificial hybrid core, which is based on BRep, the difficulty of comparing data sets given in different data models and dimensions is consequently reduced to the problem of comparing 2D or 3D points, lines, surfaces, and solids.

Third, applying the basic modeling principles of the ISO 19107 standard includes the interpretation of geometries as sets of points. For solids and polygonal objects, such point sets consist of all points that cover the objects' surfaces; for a line object, the point set comprises the line points; for a point object, the set is defined by a single point. Understanding arbitrary geometries as point sets is a further simplification because now it is not necessary to distinguish between points, lines, surfaces, and solids any more.

To give an example for entirely overlapping object representations, we refer to the situation illustrated in Fig. 3 (right). Here, a building is represented by several data sets showing significant discrepancies. The usage of the hybrid modeling concepts described above transfers the complex consistency analysis between heterogeneous data sets—originally given in different data models—to the much simpler problem of comparing structurally homogeneous point sets. These point sets, however, may appear in different dimensions. When 2D and 3D data has to be compared, the z-coordinates of the 3D point sets are neglected. Doing so, we create dimensionally adapted sets of points, as can be seen in Fig. 5 (left). For the comparison and analysis of such point sets, we can fall back on a number of approaches and metrics that have already been developed (Alt and Guibas 1996).

5.2 Hybrid Consistency for Partially Overlapping or Adjacent Object Representations

We use the term "partially overlapping or adjacent object representations" to denote data sets that mainly describe different regions of the world but, at the same time, are connected to each other due to geometric correspondences. These can be data sets showing real-world objects of the same type, such as two partially overlapping street networks from different providers. Beyond this, the object representations can also show different object types, such as an indoor model of a building on the one hand whose entrance is connected to a street network on the other hand. In both cases, the determination of hybrid consistency is restricted to those entities that are represented in both data descriptions.

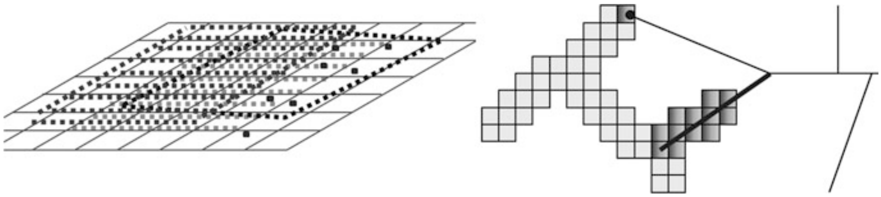

Fig. 5 Consistency analysis for entirely overlapping object representations (*left*) and consistency analysis for partially overlapping object representations (*right*)

The basis for evaluating hybrid consistency for partially overlapping or adjacent object representations is given by the possibility to explicitly model geometric correspondences as hybrid identities. The geometries to which a hybrid identity refers can be interpreted as multiple representations of local entities showing an entire overlap. Thus, the comparison of the respective geometries can be treated in the same way as discussed in Sect. 5.1.

As an example for partially overlapping data sets, we extend the geometric configurations illustrated in Fig. 4b–e to a scenario of two different network representations. As indicated in Fig. 5 (right) by the geometries highlighted in red (bold), data sets may be connected through more than one hybrid identity. One possibility to get an overall consistency value is to compute a weighted mean out of the consistency values determined for all individual hybrid identities. The weight of a hybrid identity either can be estimated from the accuracy of the geometries involved or may result from the ratio of the spatial extension of its geometries compared to the spatial extension of the geometries of all other hybrid identities. Detailed investigations on these and further possibilities for aggregating consistency values of several hybrid identities will be part of our future work.

6 Conclusions and Outlook

We proposed a data model that is meant to provide an application-independent conceptual basis for smart geoinformation systems. The data model is hybrid in the sense of structural and geometric aspects. Through targeted extensions of an existing ISO standard, our concept is able to bridge the gap between 2D, 2.5D, and 3D data and break down barriers between various modeling strategies. The consideration of geometric and topological heterogeneity is realized on the conceptual level: Hybrid identities can be defined for various objects or object parts whether they are geometrically and topologically consistent to each other or not. The explicit modeling of such geometric correspondences allows not only for the connection of objects or object parts given in different types, geometric data models, dimensions, and quality levels, but it also supports consistency analyses and updating measures, which is an important aspect considering the frequently occurring changes in geodata. The system supports multiple representations that can be based on either the same or differing data models. Additionally, it is also possible to model parts of a single object using different modeling concepts. Although, for example, the main body of a building can efficiently be represented by cell decomposition, decorative elements such as 2.5D reliefs could be added as fine surface meshes.

In future work, we will evaluate the efficiency and the potential of our hybrid modeling concept based on application scenarios. One application might be mapping and integrating multiply represented inconsistent building data into our hybrid data concept, and modeling hybrid identities for corresponding geometries. Another scenario could be the connection of disjoint or only partially overlapping data sets—for example, vector representations of street data and raster images of evacuation plans representing the interior of buildings—as basis for an outdoor–indoor navigation.

Through the integrated evaluation of geodata from different sources covering different aspects of real-world objects, we expect a deeper insight in geometric but also semantic relations. Explicitly defined hybrid identities constitute links between various data sets, and, thus, provide a basis for the inference and comprehension of higher context.

References

Alt H, Guibas LJ (1996) Discrete geometric shapes: matching, interpolation, and approximation: a survey. In: Technical report B 96-11, EVL-1996-142, Institute of Computer Science, Freie Universität Berlin

Andrae C (2009) Spatial Schema—ISO 19107 und ISO 19137 vorgestellt und erklärt. From the series. In: Andrae C, Fitzke J, Zipf A (eds) OpenGIS essentials—Die Geo-Standards von OGC und ISO im Überblick. Wichmann, Heidelberg

Bishr Y (1998) Overcoming the semantic and other barriers to GIS interoperability. Int J Geograph Inf Sci 12(4):299–314

Chen T, Schneider M (2009) Data structures and intersection algorithms for 3D spatial data types. In: Proceedings of the 17 ACM SIGSPATIAL, pp 148–157

Dakowicz M, Gold C (2010) A unified spatial model for GIS. In: IAPRSSIS, Hong Kong, vol XXVIII, part 2, pp 22–27

Fritsch D (1988) Hybride graphische Systeme—eine neue Generation von raumbezogenen Informations systemen. Geo-Informations-Systeme 1(1):12–19

Gröger G, Kolbe TH (2003) Interoperabilität in einer 3D-Geodateninfrastruktur. In: Proceedings of Münsteraner GI-Tage 2003, pp 325–343

Gröger G, Plümer L (2011) How to achieve consistency for 3D city models. GeoInformatica 15(1):137–165

Gröger G, Benner J, Dörschlag D, Dress R, Gruber U, Leinemann K, Löwner MO (2005) Das interoperable 3D-Stadtmodell der SIG 3D. ZfV 6:343–353

Herring J (2001) The OpenGIS abstract specification, topic 1: feature geometry (ISO 19107 spatial schema), version 5. OGC Doc. No. 01-101

Inhye P, Hyeyoung K, Chulmin J (2007) 2D-3D Hybrid data modeling. In: Proceedings of the 27 annual ESRI user conference, San Diego (on CD-ROM)

Kolbe TH, Gröger G, Plümer L (2005) CityGML—interoperable access to 3D city models. In: Oosterom P, Zlatanova S, Fendel EM (eds) Proceedings of the 1st international symposium on geoinformation for disaster management. Springer, pp 883–899

Lee J, Zlatanova S (2008) A 3D data model and topological analyses for emergency response in urban areas. In: Geospatial information technology for emergency response, Taylor & Francis Group, London, UK, pp 143–168

Peter M (2009) Presentation and evaluation of inconsistencies in multiply represented 3D building models. In: Rothermel K, Fritsch D, Blochinger W, Dürr F (eds) Quality of context: first international workshop, QuaCon 2009, Stuttgart, pp 156–163

Proctor MD, Gerber WJ (2004) Line-of-sight attributes for a generalized application program interface. J Defense Model Simul: Appl Methodol Technol 1(1):43–57

Stekolschik A (2007) Beitrag zum ganzheitlichen Qualitätsmanagement von CAD-Modellen in der Produktentstehung. Dissertation, Uni Bochum

Volz S, Walter V (2004) Linking different geospatial databases by explicit relations. In: Proceedings of the 20 ISPRS congress, Com IV, Istanbul, pp 152–157

Voudouris V (2010) Towards a unifying formalisation of geographic representation: the object-field model with uncertainty and semantics. Int J Geograph Inf Sci 24(12):1811–1828

Walter V (1997) Zuordnung von raumbezogenen Daten—am Beispiel von ATKIS und GDF. Dissertation, DGK, Reihe C, Nr. 480, München