Computers, Environment and Urban Systems 34 (2010) 496–507

Contents lists available at ScienceDirect

Computers, Environment and Urban Systems

journal homepage: www.elsevier.com/locate/compenvurbsys

Generating web-based 3D City Models from OpenStreetMap: The current situation in Germany

M. Over *, A. Schilling, S. Neubauer, A. Zipf

Department of Geography, University of Heidelberg, Berliner Straße 48, D-69120 Heidelberg, Germany

ARTICLE INFO

Keywords:
3D City Models
Volunteered geographic information
3D visualizations
Web 3D service
Web services
OpenStreetMap

ABSTRACT

This paper investigates the prospects for the generation of interactive 3D City Models based on free geo-data available from the OpenStreetMap (OSM) project and public domain height information provided by the Shuttle Radar Topography Mission. In particular, the suitability and quality of the OpenStreetMap data for 3D visualizations of traffic infrastructure, buildings and points of interest (POIs) is reviewed. The diversity and quantity of the points of interest provide new opportunities and challenges in creating customized and detailed visualization of cities. Specialized web services were implemented to filter and display the data in an acceptable manner. All applied web services of the 3D spatial data infrastructure are based on standards and draft specifications of the open geospatial consortium (OGC). The service is available online at www.osm-3d.org.

1. Introduction

User generated geo-information is a relatively new phenomenon in geoinformatics. OpenStreetMap (OSM) is the most prominent example of community based mapping projects. Regional differences in map coverage and accuracy make it difficult to make general statements about the suitability of OSM as a replacement for professionally captured and edited data. In some urban areas, the quantity of roads in the network already surpasses that of commercial or governmental data sets (Haklay, 2009; Haklay, under review; Zielstra & Zipf, 2009). This enables us to investigate whether OSM data can be used for creating virtual City Models.

Virtual 3D City Models can be used in different application areas such as mobile telecommunications, disaster management, urban planning, etc. (see Table 1). The applications have different demands regarding the accuracy of the 3D models and the types of 3D data needed. For example, navigation applications for mobile devices do not necessarily need a 2.5D digital terrain model (DTM), whereas this is common for car navigation applications. The following requirements for the selected application areas and target groups in Table 1 are based on qualitative interviews and questionnaires with the staff of German government departments that are concerned with geographical data. Additionally the expertise of the special interest group 3D (www.sig3d.de) members from the target groups in Table 1 were incorporated (Städtetag Nord Rhein-Westfalen, 2009).

The open geospatial consortium (OGC) standard "City Geography Markup Language (CityGML) Encoding Standard" is a common semantic information model for the representation of 3D urban objects (Geospatial Consortium (OGC), 2008). The CityGML standard also proposes accuracy requirements for different levels of detail (LODs) for virtual City Models (Table 2).

In this paper, we investigate the following issues:

1. Which CityGML standard accuracy requirements of the LOD categories in Table 2 can be achieved based on the OpenStreetMap dataset.
2. Which applications listed in Table 1 could have their accuracy needs fulfilled. Note that the 3D point accuracy needs of the applications could differ from the City GML LOD schema.
3. How could the accuracy requirements in Table 2 be achieved in the future by new OSM attributes or measurement techniques.
4. Furthermore, we investigate how the OSM data could be visualized in 3D and how OGC location based services (LBS) (OpenLS: Open Geospatial Consortium (OGC), 2005) can be used for adding value to the OSM dataset. Studies have shown that LBS are an important part in the value added chain (Fornefeld, Oefinger, & Jaenicke, 2004). The value of the geo-data to the user is not inherent in the data itself, but is added by services that select and present the data according to the user's demands.

At present, no literature exists describing the 3D visualization of OpenStreetMap data. There are only two applications that are able to render the OSM data in 3D. These OSM 3D Viewers, KOSMOS WorldFlier (Worldflier, 2009) and OSM3D (OSM3D, 2009), are very limited in the size of the scene depicted and the selected data that

* Corresponding author. Tel.: +49 0 6221x545533; fax: +49 0 6221x544996.
E-mail address: over@uni-bonn.de (M. Over).
URL: http://www.geog.uni-heidelberg.de/lehrstuehle/gis/index.html (M. Over).

0198-9715/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.compenvurbsys.2010.05.001

Table 1
Target groups and application areas for digital city models and digital terrain models (DTM) (Source: Städtetag Nord Rhein-Westfalen (2009)).

Application area	3D Data	Accuracy, DEM raster	Level of detail: (CityGML LOD)
Urban planning	DTM	DTM 1 m	LOD 2–4
	City model	City model < 0.5 m	
Traffic planning	DTM	DTM: 0.3–0.1 m	LOD 3
	City model	City model: 0.1 m	
Environmental noise mapping	DTM	0.1 m	LOD 2
	City model		
Dispersal of pollutants	DTM	5 m	LOD 1
	City model		
Security services	DTM	0.5–5 m	LOD 2–3
Disaster management	City model		
Communal business development	City model	5 m	LOD 3–4
Communal tourism development	DTM	DTM < 10 m	LOD 3–4
	City model	City model < 2 m	
Telecommunications	DTM	DTM 2–10 m	LOD 1–3
	City model	City model < 0.5 m	
Professional flight simulation	DTM	DTM: 25–50 m	LOD 0
	Texture	Texture: 1 m	
	3D airport model	3D Airport model: 0.5–1 m	LOD 2–3
Navigation (mobile clients)	City model	0.5 m	LOD 2–3
Navigation (car)	DTM	DTM: 10 m	LOD 2–3
	City model	City model: 0.5	

Table 2
LOD 0–4 of CityGML with its accuracy requirements (Source: Open Geospatial Consortium (OGC) (2008)).

	LOD 0	LOD 1	LOD 2	LOD 3	LOD 4
Model scale description	Regional landscape	City Region	City districts projects	Architectural models (outside), landmarks	Architectural models (interior)
Class of accuracy	Lowest	Low	Middle	High	Very high
Absolute 3D point accuracy (position/height)	Lower than LOD1	5/5 m	2/2 m	0.5/0.5 m	0.2/0.2 m
Generalization	Maximal generalization (classification of land-use)	Object blocks as generalised features; >6 $*$ 6 m/3 m	Objects as generalised features; >4 $*$ 4 m/2 m	Objects as real features; >2 $*$ 2 m/1 m	Constructional elements and openings are represented
Building installations	–	–	–	Representative exterior effects	Real object form
Roof form/structure	No	Flat	Roof type and orientation	Real object form	Real object form
Roof overhanging parts	–	–	n.a.	n.a.	Yes
City furniture	–	Important objects	Prototypes	Real object form	Real object form
Solitary vegetation object	–	Important objects	Prototypes, >6 m	Prototypes, >2 m	Prototypes, real object form
Plant cover	–	>50 $*$ 50 m	>5 $*$ 5 m	<LOD2	<LOD2

can be visualized in 3D. KOSMOS WorldFlier does not render buildings in 3D and OSM3D does not support a digital elevation model (DEM). Further OSM data like POIs were not visualized by either application. Literature about the automatic generation of 3D City Models from 2D data will be discussed in detail in Section 3.

The technologies and expertise regarding the creation of virtual City Models have been investigated during previous projects (GDI-3D, 2009; Schilling, Basanow, & Zipf, 2007; Zipf, Basanow, Neis, Neubauer, & Schilling, 2007). In addition, integration into an existing spatial data infrastructure using existing standards (or the proposed standards of the OGC) is of special importance to us so as to support full interoperability. Dedicated web service interfaces for supporting perspective and fully immersive 3D map applications are currently under development. The technical details are not covered in this article. The results of this project have been made publicly available through a Web 3D Service (Geospatial Consortium (OGC), 2005), which serves parts of the 3D City and Landscape Model as VRML models, and through an interactive web client (XNavigator), which allows free navigation through the complete data set and access to other OGC services (routing, geocoding and others).

2. Data Sources

In addition to the free OpenStreetMap 2D vector data, height information is also needed for 3D visualizations. As with most "commons-based peer-productions" (Benkler, 2005) the OSM data has a copyleft license, specifically the Creative Commons Attribution-Share Alike 2.0 (CC-BY-SA 2.0) license (Creative Commons Attribution-Share Alike 2.0 (CC-BY-SA 2.0), 2009). This license gives permission to any person to reproduce, adapt or distribute the work as long as any resulting copies or adaptations are also bound by the same copyleft licensing scheme. Therefore the data cannot be mixed with proprietary DEMs. To avoid licence problems we chose height data from the shuttle radar mission (SRTM) – which is in the public domain.

2.1. OpenStreetMap project

The OSM project is one of the most impressive examples of user-generated content (Benkler, 2005). Worldwide, volunteers gather geographic information as "intelligent sensors" (Goodchild,

2007) via GPS or by digitizing orthophotos or satellite data. Volunteers then make this data available via OSM. The development of the project (which started in 2004) is impressive and the quantity of the data almost tripled last year compared with the year before.

The street network is already nearly complete for some cities in Germany. Comparing the OSM data set to data from the surveying office Hamburg, it appears that the OSM data for Hamburg already covers about 99.8% of the street network (OpenStreetMap, 2009). Besides the street network, the real advantage of the dataset is the availability of manifold points of interest (POIs). POIs are point locations which may be useful or interesting. Thus the 3D City Model can be enriched with numerous items of information, making the map much more informative for the user (Fornefeld, Oefinger, & Jaenicke, 2004).

Haklay (Haklay et al., under review, 2009) shows that OSM data is superior to the official dataset for Great Britain Meridian 2; despite this, the data is not widely used in the geoinformatics. This may be the case because the usual standards for the quality of the maps and error estimation are not relevant in the case of OSM, in which there are pronounced regional differences and very rapid changes in a short period of time. Companies like Yahoo! and automotive navigation data (AND), as well as public land surveying offices in France, England, USA and Germany, all provide geo-data to the project. These companies and groups jointly profit from the data enrichment and data correction resulting from the activities of the OSM community. The following sections describe how OSM content can be used in order to create virtual City Models.

2.2. Public domain terrain data

The digital elevation model (DEM) from the shuttle radar topography mission (SRTM) is the almost global public domain dataset with the highest available spatial resolution and therefore our first choice. SRTM provides a so-called digital surface model (DSM). The difference between this DSM and a digital terrain model (DTM) is that plants and buildings are not filtered out so that the reported elevation in urban and forest areas is higher than the actual ground surface. The term DEM is used here as the generic term for DSMs and DTMs, only representing height information without any further definition about the surface. This is the most common usage in literature where varied definitions for the terms DEM and DTM exit, an overview is given by Li, Zhu, and Gold (2005).

Multiple SRTM versions, produced using different processing algorithms or additional height information, are available. Due to the side-looking radar technique used by SRTM, distortions may be visible and voids occur where no data could be recorded (Henderson & Lewis, 1998).

CGIAR (2009) provides (in version 4) a corrected SRTM dataset in which those voids have been compensated for by using additional sources. The spatial resolution is three arc-seconds (about 90 m). The absolute vertical error for Germany is up to +/−7 m depending on the surface structure (Czegka, Behrends, & Braune, 2004). The CGIAR dataset is not public domain, but since it is not possible to reconstruct the original dataset from the integrated DEM, which is generated by our process, CGIAR provided us with their dataset and allowed us to licence it under the CC-BY-SA 2.0 licence.

3. Architecture

This section presents a short overview of the system architecture. Detailed information has been presented in other publications (Schilling et al., 2009; Zipf, Basanow, Neis, Neubauer, & Schilling, 2007). In the former project, "Spatial Data Infrastructure 3D" GDI-3D (2009), one of the first implementations of the Web 3D Service (Open Geospatial Consortium (OGC), 2005) based on an OGC draft specification was developed. The result from a W3DS getScene-Request is a 3D scene-graph in a common 3D file format (e.g. VRML, X3D.). We chose the VRML format because it is widely supported by 3D browser plug-ins and the compressed VRML encoding is more effective than other compressed encodings like X3D, 3ds and Viewpoint (Schilling, Basanow & Zipf, 2007). The 3D scene could be visualized using conventional browsers with a 3D plug in. Visualization with our own 3D Viewer, the XNavigator, is much more user friendly because the desired requests depending on the location of the user are sent automatically to the W3DS. Layers can be individually selected by the user, and further applications (routing, searching for addresses and POIs) based on OGC Web Services are implemented. All layers could be styled by the user on demand via a 3D styled layer descriptor (3D-SLD), which is an enhancement of the OGC SLD standard (Neubauer & Zipf, 2007).

4. Combining land-use areas with terrain

In principle, there are three alternative ways to display OSM land-use data in 3D. This data can be displayed by mapping raster images onto a digital elevation model (DEM), by overlaying vector data on the DEM or by combining the vector data and the DEM in an integrated triangulated irregular network (TIN).

4.1. Texture wrapping

Direct wrapping of the textures over a DEM is commonly performed in virtual globes. Haeberling (1999) introduced a static approach for combining vector and raster layers into a single texture which is wrapped over the DEM. A dynamic method has also been

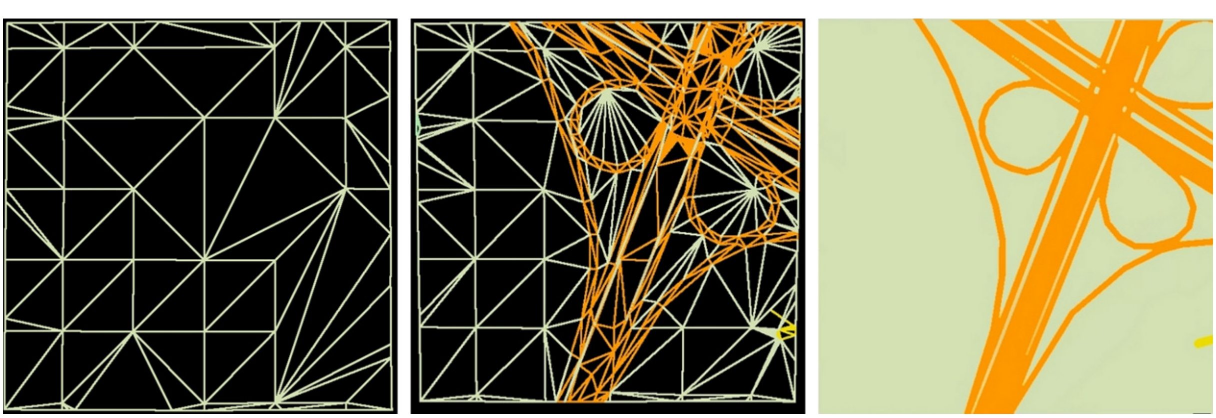

Fig. 1. Process of integrating roads into the DEM. Left: original DEM triangulation, middle: DEM with highway, right: same, with Gouraud shading.

developed by Kersting & Döllner (2002) where only the vector data is streamed over the network. The texture pyramids are rendered on demand by an off-screen pixel buffer. Dynamic rendering with an off-screen pixel buffer is not supported by standard web browsers and it is therefore incompatible with the current OGC standards.

The advantage of the texture wrapping method is that the TIN that was derived from the DEM only needs to be computed once. OSM data could be automatically extracted from a web map service (WMS) which could easily be kept up-to-date. On the downside, the requirements on the client side are very high (e.g., file transfer, graphics card, memory consumption). In general, when using high-resolution orthophotos for this purpose, distortions (e.g. slanted buildings and bridges) and aliasing effects may occur at close range.

4.2. Vector overlay

Wartell et al. (2003), Agrawal, Radhakrishna, and Joshi (2006) and Schneider, Guthe, and Klein (2005) introduced methods to overlay the vector data on top of the DEM as a separate layer. This is an improvement compared to the texture wrapping methods, alleviating the system requirements on the client side and reducing aliasing. A disadvantage is that this causes a huge overhead, especially for large polygons, because it doubles the number of triangles to render (one set for the DEM, and a duplicate set for the overlaying land-use).

4.3. Integrated DEM

Schilling et al. (2007) avoided this overhead problem by using the vector overlay method-taking the coordinates of the OSM polygons into account as new points within the triangulation (Fig. 1).

The resulting triangles within the polygon receive the attributes of the source features and can be coloured for visualization. Another advantage of this approach is that all layers could be styled by the user on demand via a 3D styled layer descriptor (3D-SLD), which is an enhancement of the OGC SLD standard (Neubauer & Zipf, 2007).

Fig. 2 presents a direct visual comparison of the texture wrapping and integrated DEM methods for a subset of 1×1 km^2. The file size of the DEM (before the integration of the vector data) was 7 KB, and after the integration, 159 KB. The JPEG image (which was wrapped over the original DEM) has a file size of 152 KB, a resolution of $2000 * 2000$ pixels and a compression ratio of 50%. Although having the same input size, the visualization of the integrated DEM is superior at close range due to the aliasing and compression effects of the JPEG image.

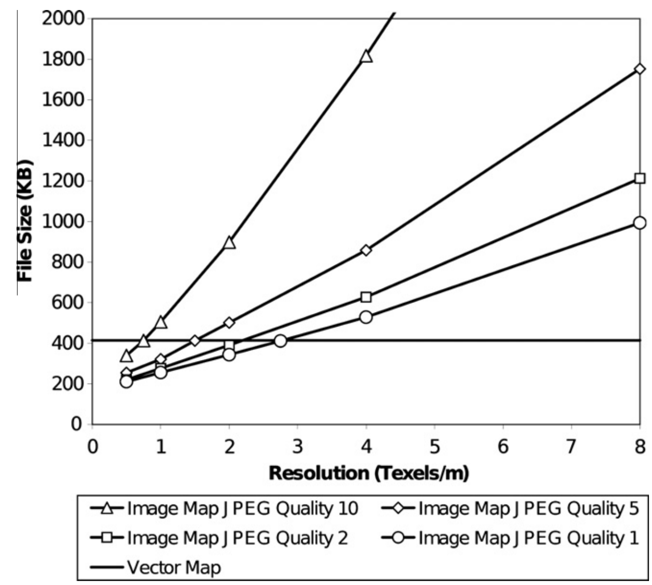

Fig. 3. File sizes of texture maps in relation to image resolution and quality, compared to the equivalent vector map (Source: Schilling et al. (2007)).

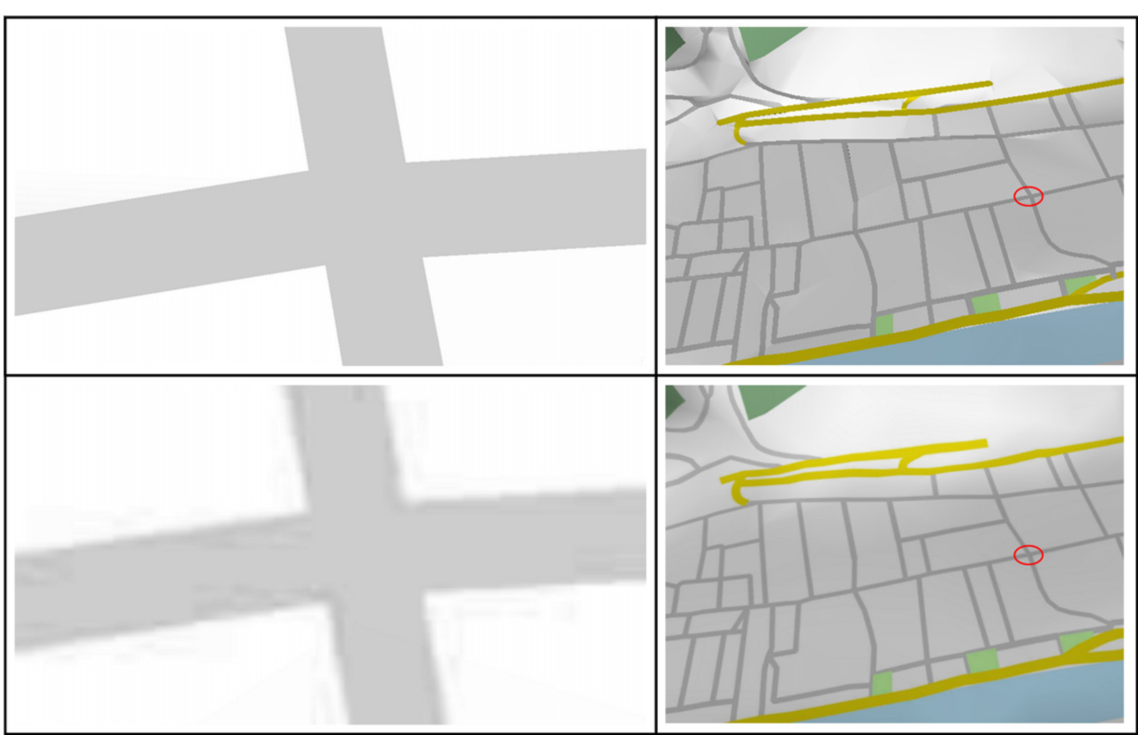

Fig. 2. Comparison of the vector map (above) and the texture map (below). The subset on the left is marked in the overview map (right).

Table 3
Applied land-use for each LOD.

Tiles / OSM Data	Sea	Water	River	Canal	Stream	Forest	Scrub	Park	Industrial	Commercial	Retail	Residential	Tertiary_ link	Secondary_ link	Primary_ link	Trunk_ link	Motorway_ link	Footway	Cycleway	Service	Track	Unclassified Roads	Living_ Street	Pedestrian	Residential	Railway	Tertiary	Secondary	Primary	Trunk	Motorway
500	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
1000	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x							x	x	x	x	x	x	x	x
2000	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x								x	x	x	x	x	x	x
4000	x	x	x	x	x	x	x	x	x	x	x	x		x	x	x	x									x		x	x	x	x
8000	x	x	x	x		x	x	x	x	x	x	x			x	x	x									x			x	x	x
16,000	x	x	x	x		x	x	x	x	x	x	x				x	x													x	x
32,000	x	x	x	x		x	x	x	x	x	x	x				x	x													x	x
64,000	x	x	x			x			x	x	x	x				x	x													x	x

Beside the visual advantages, the file size of the vector map is smaller for higher resolutions. In Fig. 3, the dependency of file sizes and resolutions are presented for JPEG images with different compression ratios and the corresponding vector map. Because the visualization of City Model usually takes place at higher resolutions, the integrated DEM method was our first choice.

The integrated DEM method requires intensive pre-processing but performance is superior for higher resolutions in the final application. This is because less data has to be streamed, in direct comparison to texture wrapping. The improved quality also results from vector overlaying. The complex pre-processing needed when using this method is described in the next section.

To add a width to the line data (streets, railways, etc.), a buffer value with mean values for each type was computed (Table 3). For lower scales, the buffer values were gradually broadened for an adapted visualization. To archive a proper draw ordering of all street types, the levels of bridges also have to be taken into account. Bridge segments in OSM contain a layer tag with values ranging from 1, for the first bridge level, up to five for the top level, which is sometimes used for very complex highway crossroads. This enables us to draw the street network in the correct order.

5. Processing of the integrated DEM

To achieve acceptable performance in the final application, various techniques were used to reduce the amount of data transferred without the loss of important information. Generally this goal was achieved by a level of detail (LOD) schema and DEM simplification. Fig. 4 shows an overview of the complex pre-processing workflow, which is explained in detail in the following sections.

5.1. LOD schema

The integrated TIN had to be computed at different levels of detail (LODs). For each LOD the complexity of the integrated DEM is decreasing. Depending on the viewpoint of the user different LODs are visualized in the W3DS client (Fig. 5).

For pre-processing, the area of Germany was divided into 42 tiles. For each tile, eight LODs (from 500 m to 64 km) were computed. The following four approaches were used to reduce the amount of data (for values see Table 4).

1. *Layer selection*: Manual selection of the required OSM features for each LOD (Table 3).
2. *Layer generalization*: A Douglas–Peucker (Douglas & Peucker, 1973) line thinning algorithm (generalization threshold) was applied on the selected LODs.
3. *Mesh simplification algorithm*: An edge contraction algorithm was applied which reduces the meshes of the TIN depending on their visual significance until a quadric error value has been reached (Garland & Heckbert, 1997).
4. *Edge factor*: Factor in order to preserve the boundaries of the 2D layers within the mesh simplification algorithm. Perpendicular planes through the edge of the polygon boundaries were computed. These planes are weighted by a large penalty factor and taken into account by the simplification algorithm (Garland & Heckbert, 1997).

The low resolution of the DEM causes skewed roads on hillsides. To avoid this problem, an adaptive DEM modification was applied (Schilling, Lanig, Neis, & Zipf, 2008). The roads are flattened using a penalty factor. This factor could be applied to each individual road type. A feature generalization was applied for the feature types of the low resolution LODs (see Table 4). For this purpose, a Douglas and Peucker (1973) line thinning approach was used.

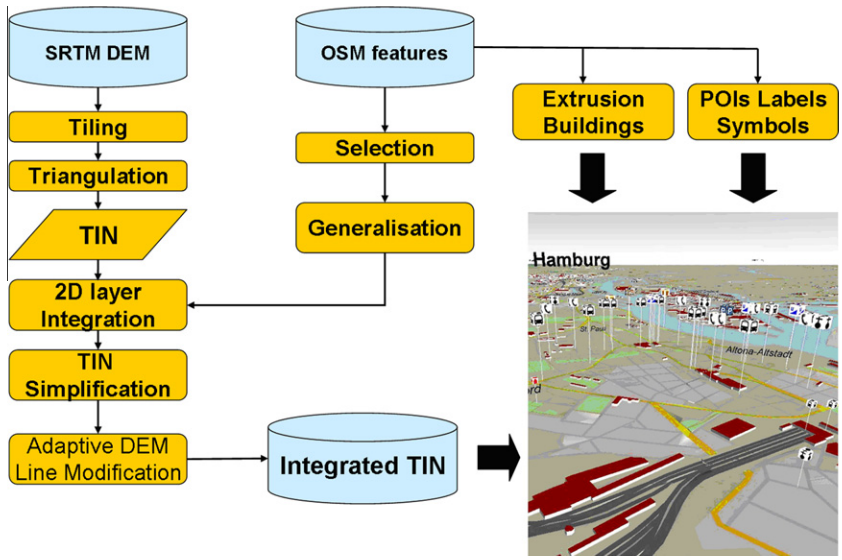

Fig. 4. Overview of the integrated DEM processing workflow.

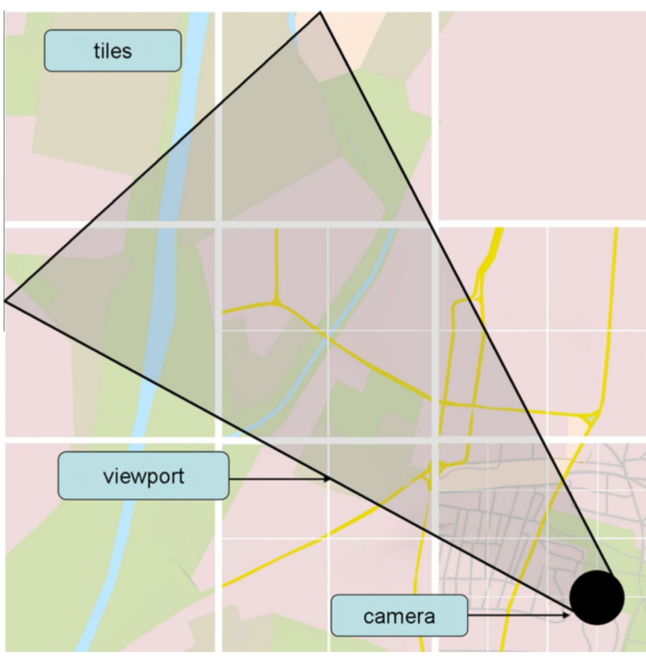

Fig. 5. Visualization of the LODs depending on the viewpoint.

Table 4
Overview of the LODs and data reduction factors.

LOD tile size in metres	2D OSM features	Simplification error	Edge factor	Generalization threshold
500	105	1	4	–
1000	69	4	4	–
2000	63	16	4	–
4000	57	64	4	–
8000	45	256	4	–
16,000	10	16,384	100	50
32,000	10	65,536	90	150
64,000	6	56,36,096	2500	350

For every LOD, tile and land-use, a VRML file was computed on a high performance cluster which resulted in a total processing time of about 1300 CPU hours and nearly 6.7 million processed VRML files.

5.2. Simplification

The simplification of the TIN was done by a modified implementation of the edge contraction algorithm of Garland and Heckbert (1997). With regard to the huge data size that had to be processed, this algorithm is an acceptable compromise between very fast (low-quality) methods such as vertex clustering (Rossignac & Borrel, 1993) and very slow (high-quality) methods such as mesh optimization (Hoppe, 1996). The implementation is described in detail in the publications of Schilling et al. (2007, 2008). The iterative algorithm reduces the meshes of the TIN depending on their visual significance until a quadric error value has been reached.

Using a semi-automatic approach, the error value for each LOD was estimated empirically under the condition that no resulting VRML tile is larger than 500 KB. On Notebook or Desktop PC's with a dedicated graphics card, 1.5 GB RAM, a Dual Core CPU and an Internet connection of 4000 kbit/s or more, the performance of the final application is judder free under the condition that the maximum tile size do not exceed 500 KB. The above-mentioned CPU and Internet requirements were a compromise, taking into account the broadness of the user group, the richness of the details and the hardware performance. A similar error value was applied to all tiles of an LOD.

The standard edge contraction approach destroys the boundaries of the polygons. To avoid that problem, a large penalty factor (edge factor) was implemented (Garland & Heckbert, 1997) in order to preserve the boundaries of the 2D layers.

6. Processing of buildings and point features

The OSM project offers a very wide range of data and the standard feature set is constantly being extended. New tags are established by a voting process. Only approved tags are recognized by the OSM renderers Mapnik and Osmarenderer and displayed in the image tiles.

All other OSM data (except land-use areas, street and rail networks, which is part of the integrated DEM), is processed individually and updated on a daily or weekly schedule. We defined a set of layers in three different categories: labels, buildings, and points of interest (POIs) (Table 5). Each layer can be later activated

Table 5
Volume of daily or weekly updated OSM features (July 2009).

Labels	Buildings	Points of interest
Region labels (200)	Buildings (416,713)	Accommodation (9652)
County labels (53)	Building numbers (336,411)	Eating (49,440)
City labels (2318)	Building labels (38,174)	Education (20,889)
Suburb labels (7175)	TechBuildings (6728)	Enjoyment (2497)
Village labels (51,027)		Health (10,945)
Locality labels (2575)		Money (13,641)
Natural labels (4213)		Post (22,369)
Street labels (1403,230)		Public facilities (58,869)
		Public transport (98,835)
		Shop (32,591)
		Traffic (96,345)

and controlled individually in the 3D client. The number in brackets in Table 5 is the feature count for Germany as of July 2009.

All features in these layers are represented as 3D geometries and placed on the terrain model. The height information for the reference point is provided by a specialized elevation query service (EQS), which plays an important role in all update processes and also in the live system. Since the complete DEM is almost impossible to load in memory, the task of the interpolation of elevation values had to be outsourced to a dedicated service.

6.1. Buildings

OSM buildings are modelled as closed rings. For creating 3D shapes, the height must be derived from other OSM attributes (called tags), or, if none is available, set to a default value. Users have also proposed advanced building attributes, which can be used to improve the visual appearance of the City Model (Table 6).

The number of OSM buildings for Germany is 408,594 (as of July, 2009). This is about 2.5% of the total number of residential buildings in Germany (Statistisches Bundesamt, 2009).

For the generation of CityGML LOD 1–4, height information for buildings is needed. At the beginning of August 2009, height information for about 2000 buildings was available in Germany. The height is tagged in metres or as the floor number of the highest floor. A direct measurement of the height is impractical for OSM mappers in many cases because of the absence of an adequate technique. The mapping of building floors is more practicable. For the 3D reconstruction, we use the absolute height value, or, if this is not available, the number of floors from which we estimate the building height. For all other buildings, we use a default value. To distinguish which kind of height information was applied to a building, different colours were assigned at visualization.

Extrapolating the current development, we predict the completion of the building footprints in Germany within the next four years. Capturing height information and other attributes will take much longer, since it cannot be taken from orthophotos or satellite images, which are usually the basis for digitizing the footprints. An alternative could be a cooperation with public authorities, who have previously donated material to the OSM project. As an example, the city of Rostock supplied their complete building data set. The building heights of 17,000 buildings (which were also included), were imported into OSM at the end of August 2009. Therefore, for the city of Rostock the first LOD 1 CityGML Model based on OSM data can be generated.

The building update workflow, which is performed daily, is shown in Fig. 6.

Attributes that may be relevant for 3D visualizations, (like the roof shape, gutter height and roof orientation) were already suggested to the OSM community by our research group. These attributes are currently only marginally tagged by the community.

Table 6
Proposed building attributes (* proposals of our research group) (OpenStreetMap (OSM), 2009; OpenStreetMap Wiki, 2009).

Key	Description	Example value
Height	Approximate height	541 m
Building:roof	Roofing material	Slate
Building:cladding		Glass (skyscrapers)
Building:type	General "type" or design of building	Skyscraper
Building:architecture	Architectural style	Victorian
Building:levels	Number of stories of the building	50
Building:use	Main use of the building	Office
Building:shape	Approximate shape of the building	Tower
Building:model	URL of 3D model of the building	http://somehost.com/building32.2dm
Building:roof:shape	Shape of the roof	Flat
Building:roof:eaves*	Eaves height of the building	15 m
Building:roof:colour*	Predominant colour of the walls	Grey
Building:colour*	Predominant colour of the roof	Red
Building:roof:orientation*	Orientation of the roof	Along
Building:levels:height*	Mean height of the levels	3 m

6.2. Points of interest

Points of interest (POIs) have been captured for an abundance of different locations, shops, restaurants, facilities, technical installations, and so on. They provide in part very deep information – which enables applications that go far beyond the static display of map content. For some categories, a tagging schema has been established for storing typically useful information about a specific type of facility. The schema for restaurants, for instance, includes name, address, opening hours, cuisine, telephone number, and URL of the homepage.

The primary OSM key for this kind of node is "amenity". The value describes the type which can be used to assign an icon or symbol. The generic "name" key may be used for an additional label. The amenity types have been divided into the categories: accommodation, eating, education, enjoyment, health, money, post, public facilities, public transport, shop, and traffic. Each category is provided as individual layer through the Web 3D Service.

The processing workflow is similar to the one for buildings. The elevation value for each POI is retrieved from the EQS. The 3D coordinate is used as a reference point for a 3D symbol, modelled as a textured box showing an image of the amenity type (Fig. 7).

Furthermore, POIs can be used for spatial searches. This functionality is provided by an OpenLS Directory Service (Open Geospatial Consortium (OGC), 2005), which is also used by www.openrouteservice.org. After clicking on a location, and specifying the POI type and a maximum distance, a request is passed to the Directory Service, which then searches its database for POIs matching the criteria. Fig. 8 shows the result of a proximity search for hotels around a start point. If more information on the POIs becomes available, like contact points, exact addresses, or reviews, then the Directory Service could be extended to become an alternative to the Yellow[1] Pages.

[1] For interpretation of color in Fig. 8, the reader is referred to the web version of this article.

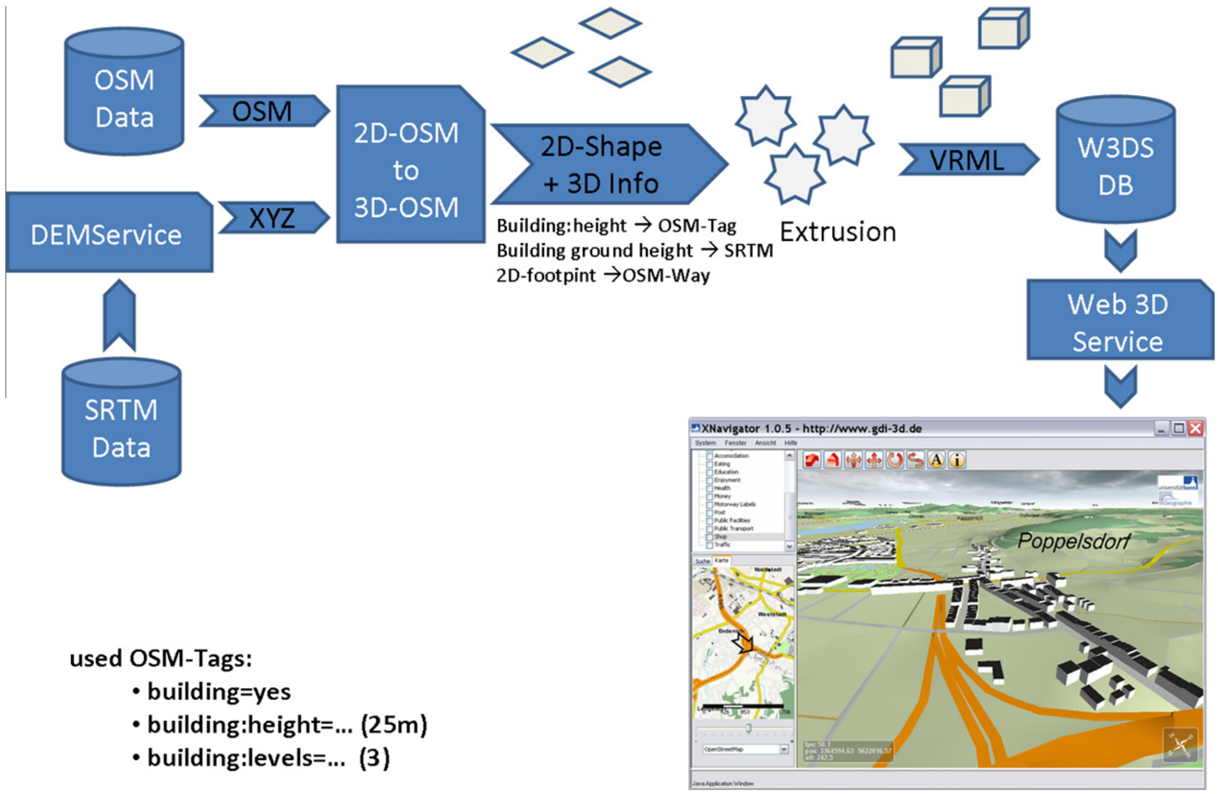

Fig. 6. 3D building creation workflow from OSM data.

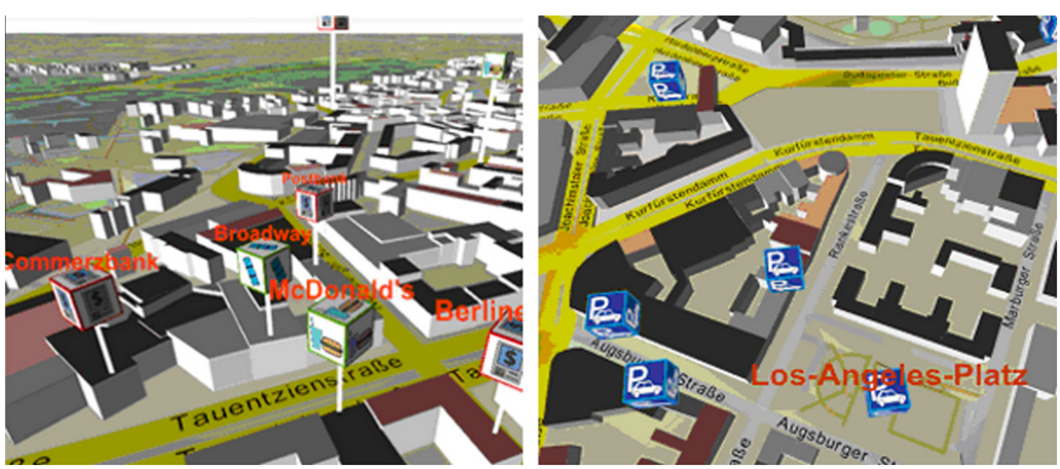

Fig. 7. Points of Interest visualized as 3D symbols and labels.

Technical facilities like wind turbines, lighthouses, wind-mills and traffic lights were also visualized as 3D models (Fig. 9), supporting the visual orientation of the user (Schilling et al., 2008).

7. Quality assessment of the OSM dataset

OSM was designed for creating user generated free maps. A multitude of online applications show the potential to do much more than just rendering maps. Height information is partly available, which enables us to build virtual City Models. The strength of OSM is the vast number of volunteers contributing to this project. However, accuracy is very limited, due to the data capture methods. Usually GPS tracks are used to record the road network, which have an inherent horizontal accuracy of about 10 m. However Haklay (2009) showed that the position accuracy of the road network of the OSM dataset is superior compared to the official Ordnance Survey dataset of Great Britain Meridian 2. Our own experiments showed that the vertical accuracy of over 1000 GPS track points was about +/−25 m compared to a surveying office DTM. Hence is below the accuracy of the SRTM DEM, and consequently cannot be used to improve it.

Building footprints are mostly digitized from aerial or satellite imagery. Legal agreements with online map providers allow mappers to use some high-resolution imagery for deriving vector data. In this case, the accuracy of the digitized building footprints depends on the resolution of the provided orthophotos and satellite images.

To estimate the accuracy of the OSM building footprints, they were compared to the Federal cadastre data called ALK (which

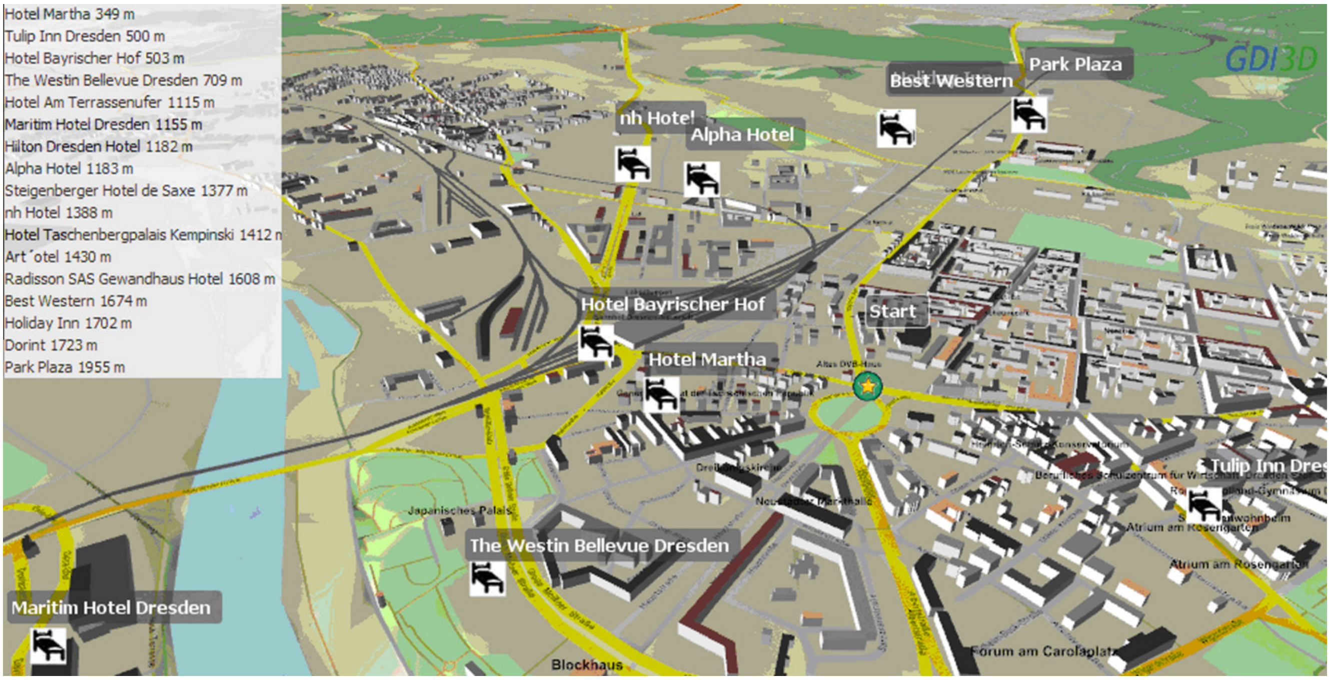

Fig. 8. Result of a proximity search for hotels using the Directory Service from www.openrouteservice.org.

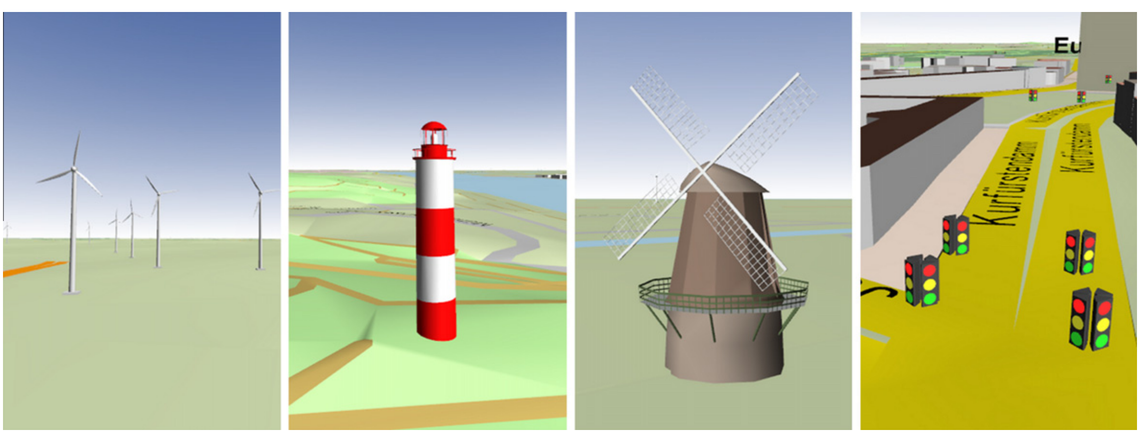

Fig. 9. Some examples of technical facilities. From left to right: wind turbines, lighthouse, traditional windmill, traffic lights.

has an accuracy of less than 0.5 m). Buffers of different sizes were computed around the ALK footprints; then the percentage of the OSM buildings within these buffers was calculated.

An absolute 3D point accuracy for the building footprints of 5 m, postulated by the CityGML LOD1, could not be achieved using the OSM dataset at this point. Currently the accuracy requirements of CityGML LOD 1 could be reached for the city of Rostock, using data donated by the local authority (see Section 6.1).

Using the number of floor levels to estimate the building heights is not a reliable method if the mean height of a building level is not tagged, which is the usual case.

Beside the positional accuracy, the completeness of the map is an important issue as regards data quality. Zielstra and Zipf (2009) compared the length of the OSM road network with the commercial dataset of TeleAtlas Germany.

As shown in Fig. 10, the completeness of the OSM dataset decreases with the distance to the city centres.

Another limiting factor is the absence of exact road shapes. The road width has to be derived from the road type. Each road type is assigned a fixed width. This is acceptable for small-scale visualizations, resulting in a map-like rendering. Virtual Reality or other 3D applications would benefit from more information about the number of lanes, road width, or exact locations of kerbstones, pavements, and road boundaries.

The quality control of OSM differs fundamentally from professionally edited maps. The community-based approach allows anyone to upload and alter map data. But due to the huge number of editors, errors and conflicts are usually quickly resolved. OSM has probably the most up-to-date map data. In urban areas, changes in the road network appear in the OSM data set long before appearing in other map providers' data (LandkartenBlog.de, 2009). The number of POIs is increasing rapidly, but is still at a low level compared to commercial providers. 21,574 bus stations and tram station out of a total of about 50,000 in North Rhine-Westphalia are currently

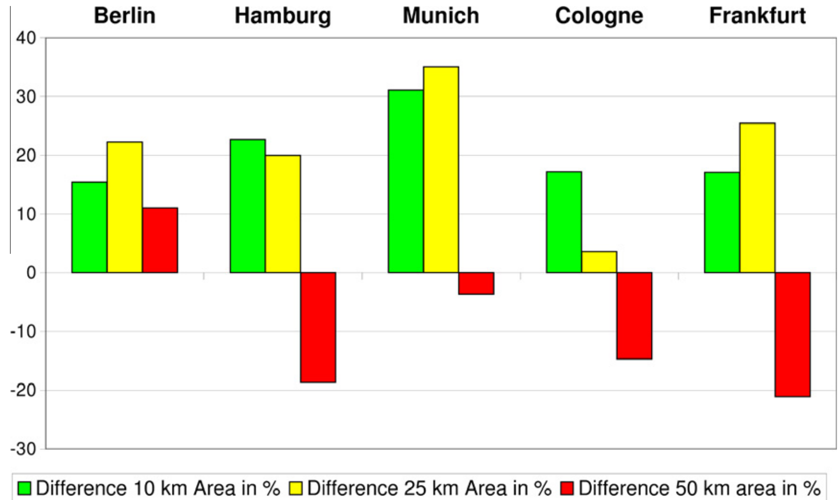

Fig. 10. Differences in the road network data as between that provided by OSM and TeleAtlas data for five major cities and the urban hinterland. Positive values represent a longer road network of the OSM dataset (Source: Zielstra and Zipf (2009).

tagged (as of April 12, 2009). The OSM community has yet to prove that it can keep track of the ever-changing "landscape" of restaurants, cafes, shops, etc.

8. Results and discussion

In this section the questions formulated in the introduction should be answered in the given order.

8.1. Applicability to demands of the CityGML LODs

First we attempted to determine whether the OpenStreetMap dataset could attain the levels of accuracy proposed in the CityGML standard for City Models (Table 2).

There is no absolute 3D point accuracy given for CityGML LOD 0. Using current measurement techniques, OSM point accuracies of about 10 m are possible.

Regarding the completeness aspect of the data quality, there is a big difference between the completeness of the urban and rural areas. The OSM's data set is significantly more complete in urban areas. This reflects the concentration of OSM mappers in highly developed urban areas (Haklay, 2009). As shown in Fig. 10, the completeness of the OSM road network surpasses that of the commercial TeleAtlas dataset for all five cities inside a buffer zone of 25 km around the cities limits. We can assume that for the German metropolitan areas, CityGML LOD 0 is available with a point accuracy of 10 m.

The requirements of CityGML LOD 1 could already be achieved for the city of Rostock (see Section 6.1). This should not hide the fact that with the currently available techniques of digitizing footprints from satellite images or orthophotos the requirements of an absolute point accuracy below 5 m (Table 2) could not be fulfilled.

CityGML LOD 2 must contain roof forms and structures. Both could be described with the proposed attributes listed in Table 6. The relative orientation of the roof (along or across the building) could be mapped by attributes or by sketching the roof lines from orthophotos and satellite images. The building height (roof ridge) and eaves height could be estimated by the user. But the investigation of the absolute point accuracy for the building footprints in Table 7 showed that an accuracy below 4 m could not achieved by digitizing techniques on the basis of the images provided.

Starting from LOD 3, all models must show the exact shape of building structures, even of facades and smaller roof elements like dormers, with an absolute point accuracy of 0.2–0.5 m. It is hardly possible to do this based on the building footprints and attributes.

Table 7
Percentage of the OSM buildings within different buffers around the Federal cadastre ALK) building footprints for the city of Cologne.

Buffer around ALK building footprint (m)	Percentage of the contained OSM buildings (%)
2.5	7
5	55
7.5	81.8
10	100

8.2. Applicability to demands of the applications and target groups

Now we attempted to determine whether the OpenStreetMap dataset could attain the demands of the applications listed in Table 1. At that point, the OSM land-use data could be used regionally for applications like car navigation, telecommunications, tourism development and, potentially, professional flight simulation. As shown above the absolute 3D point accuracy below 10 m is available for the German metropolitan areas. Beside the landscape model the applications need city models with an absolute 3D point accuracy which could not be attained on base of the OSM dataset yet. Only the city model of Rostock attains the 3D point accuracy requirements for dispersal of pollutants applications (see Section 6.1).

Regarding the requirements of the digital terrain model accuracy for the applications listed in Table 1, it is clear that most of these applications cannot be based on SRTM data with an absolute accuracy of +/−7 m. The SRTM mission provides a digital surface model, but for the above-mentioned applications a digital terrain model is needed. To match the requirements of the applications, higher resolution digital terrain models must be made available in the public domain or at least with an OSM-compatible license.

8.3. Improvement through advanced OSM attributes and measurement techniques

The absolute point accuracy could be increased using higher resolution images in the future. Beside the Company Yahoo! more and more surveying offices in Germany provide high resolution images for the OSM project. Also advanced GPS augmentation techniques like EGNOS (www.esa.int/esaNA/egnos.html) could improve the absolute accuracy of the OSM dataset in the future.

The computation of building heights using the number of floors alone is not a reliable method to measure the height of a building. A mean floor height attribute would be necessary to achieve an absolute point accuracy below 5 m (LOD 1, Table 2). Cooperation with public authorities would be the fastest and most promising way to achieve CityGML LOD 1.

Another issue is complex buildings with differing roof levels. In this case the building could be divided into different parts. To allow, for example, the calculation of the roof heights of a building (e.g. for urban planning applications), new OSM attributes must be introduced that ensure that all relevant parts belong to one specific building.

LOD 3 and LOD 4 models must show the exact shape of building structures. Therefore architectural 3D models are needed. A major step forward would be the development of a free 3D Virtual City Model Database. Numerous 3D models (that can be loaded in Google Earth) have already been created by volunteers (3D data warehouse, 2009). There is already the OpenSceneryX project (OpenSceneryX project, 2009), which is an open library of 3D feature symbols with over 300,000 3D objects under a Creative Commons license. But an open project for real world architectural 3D models, which also takes the quality of the 3D models into account, is still lacking.

8.4. Visualization of OSM data in 3D

As shown in Sections 4 and 5 the OSM land-use data was integrated into the digital elevation model. This approach results in an improved visualization and data reduction for higher resolutions, compared to texture wrapping (see Section 4.1). The advantage of this integration, compared to simple vector overlaying, is that no doubled triangles for the land-use polygons and the DEM occur, cutting the memory requirement in half (see Section 4.2). The disadvantage of integrating this data is the time consuming pre-processing. Depending on the amount and type of the vector data, the integration typically takes 10 times longer than the initial terrain triangulation. Contrary to texture wrapping or vector overlay the integrated DEM must be processed for every update. The pre-processing was done in this case on a high performance cluster and resulted in a total processing time of about 1300 CPU hours.

In Section 6 the workflow for the generation of 3D Buildings and POIs was introduced. The OpenStreetMap project offers a unique database of POIs, which could be tailored on demand by an OpenLS Directory Service to fit to the needs of the user (see Section 6.2). In this way, individual interactive 3D maps could be provided. Other OGC Open Location Services (which could not be referred to in this paper) have also been successfully implemented for further applications like routing and geocoding, to add further value to the OSM dataset (Neis, Schilling, Zipf, & 2007; Neis, Zipf, & Schmitz, 2008; Schilling et al., 2009).

9. Conclusion

The huge amount of data, rapid development, promising quality analysis and the increasing data supplements from governmental and commercial map agencies have enabled us to investigate whether OSM data can be used for creating virtual City Models. For most applications listed in Table 1, and in most places that we considered, the OSM dataset does not at present fulfil the demands regarding completeness and absolute accuracy.

Nevertheless OpenStreetMap has a huge potential for fulfilling the requirements of CityGML LOD 0–1 regionally in the next few years if the development of the OSM project continues as it has in the past. For metropolitan areas, LOD 0 is already available with a point accuracy of 10 m.

Whether LOD 1 would be regionally available in the near future depends on the awareness of the community of the importance of City Models. Beside that, cooperation with federal, state and local authorities would be the fastest and most promising way to achieve that goal.

For the creation of CityGML LOD 2 from OSM, advanced OSM tags like roof type, eaves height and roof orientation would be necessary. These attributes have already been proposed to the community by our research group. Detailed architectural models (which are necessary for many applications) could not be provided by the OSM project. The setup of an architectural model database would be required for this goal.

The suitability of the OSM data for specific purposes must be carefully investigated, due to the pronounced regional differences in the quality of the dataset.

The introduced concepts and services developed in the GDI-3D project have been successfully applied to visualize the OSM data in 3D for the entire area of Germany. Location based services like routing, spatial searches and address searches were used to add further value to the dataset. The system has proved to be stable and in further projects like North Rhine-Westphalia 3D (NRW-3D) North Rhine-Westphalia 3D, 2009) and GDI-3D, it was confirmed that the 3D spatial data infrastructure is even able to handle millions of buildings and high resolution DTMs under the recently introduced LOD schema. It was shown that a 3D spatial data infrastructure could be based on the existing standards or the proposed standards of the OGC, ensuring full interoperability.

References

Agrawal, A., Radhakrishna, M. & Joshi, R. C. (2006). Geometry-based mapping and rendering of vector data over LOD phototextured 3D terrain models. WSCG 2006 – The 14th international conference in central Europe on computer graphics. Visualization and computer vision. January 30–February 3. Plzen – Bory, Czech Republic.

Benkler, Yochai (2005). coase's penguin, or, linux and the nature of the firm. In: Ghosh, Rishab A. (Hg.): CODE. Collaborative ownership and the digital economy. Cambridge, MA: MIT Press, pp. 169–206.

CGIAR-CSI (2009). Consultative group on international agricultural research – Consortium of spatial information. <http://srtm.csi.cgiar.org/>. 10.12.2009.

Creative Commons Attribution-Share Alike 2.0 (CC-BY-SA 2.0) Licence (2009). <http://creativecommons.org/licenses/by-sa/2.0/>. 10.12.2009.

Czegka, W., Behrends, K. & Braune, S. (2004). Die Qualität der SRTM-90m Höhendaten und ihre Verwendbarkeit in GIS. 24. Wissenschaftlich-Technische Tagung der DGPF. 15.-17.9.2004. Halle.

Douglas, D., Peucker T. (1973) Algorithms for the reduction of the number of points required to represent a digitised line or its caricature. In: The Canadian Cartographer, Vol. 10, pp. 112–122.

Fornefeld, M., Oefinger, P. & Jaenicke, K. (2004). Nutzen von Geodateninfrastrukturen. <http://www.micus.de/pdf/micus_gdi_studie_ 12_10_2004.pdf>.

Garland, M. & Heckbert, P. (1997). Surface simplification using quadric error metrics. In: Computer Graphics (SIGGRAPH '97) Proceedings, pp. 209–216.

GDI-3D (2009). Spatial Data Infrastructure for 3D-Geodata. <http://www.gdi-3d.de>. 10.12.2009

Goodchild, M. F. (2007). Citizens as sensors: web 2.0 and the volunteering of geographic information. In: GeoFocus (Editoral), Vol. 7, pp. 8–10.

Google 3D Data Warehouse (2009). <http://sketchup.google.com/3dwarehouse/>. 10.12.2009.

Haeberling, C. (1999). Symbolization in topographic 3D Maps. conceptual aspects for user-oriented design. In: Proceedings of 19th international cartographic conference ICA '99, Ottawa, Canada, pp. 1037–1044.

Haklay, M. (under review). How good is OpenStreetMap information? A comparative study of OpenStreetMap and ordnance survey datasets for London and the rest of England. Environment & Planning.

Haklay, M. (2009). Beyond good enough? Spatial data quality and OpenStreetMap data. In: State of the map conference 2009, Amsterdam. <http://www.slideshare.net/mukih/beyond-good-enough-spatial-data-quality-and-openstreetmap-data>. 10.12.2009

Henderson, F. M., Lewis, A. J. (1998). Principles and applications of imaging radar, 3rd ed, Vol. 2. New York: Wiley. Manual of Remote Sensing.

Hoppe, H. (1996). Progressive meshes. In: SIGGRAPH '96 Proc., August 1996, pp. 99–108.

Kersting, O. & Döllner, J. (2002): Interactive 3D visualization of vector data in GIS. In: Proceedings of the 10th acm international symposium on advances in geographic information systems – ACMGIS, November 2002, Washington, DC, pp. 107–112.

KOSMOS Worldflier (2009). I. Brejc. OSM 3D Viewer. <http://igorbrejc.net/category/3d>. 10.12.2009.

LandkartenBlog.de (2009). Der große Landkartentest. <http://landkartenindex.blogspot.com/2009/02/landkartentest-der-grote-landkartentest.html>. 10.12.2009.

Li, Z., Zhu, Q., & Gold, C. (2005). *Digital terrain modeling: principles and methodology*. Boca Ranton: CRC Press. pp. 7–9.

Neis, P., A. Schilling, A. Zipf (2007). 3D emergency route service (3D-ERS) based on OpenLS specifications. GI4DM 2007. In: 3rd international symposium on geoinformation for disaster management, Toronto, Canada.

Neis, P., Zipf, A, Schmitz, S. (2008). OpenRouteService.org – combining open standards and open geodata. In: The state of the map. 2nd open street maps conference, Limerik, Ireland.

Neubauer, S., Zipf, A. (2007). Suggestions for extending the OGC styled layer descriptor (SLD) specification into 3D – towards visualisation rules for 3d city model. In: Urban data management symposium, UDMS 2007, Stuttgart, Germany.

NRW3D – North Rhine Westphalia 3D (2009). <http://www.nrw-3d.de/>. 10.12.2009.

OpenLS: Open Geospatial Consortium (OGC) (2005). OpenGIS location service implementation specification: Core services. Ref. No.OGC 05-016.

Open Geospatial Consortium (OGC) (2008). OpenGIS city geographic mark language (CityGML) encoding standard. OGC Ref No. OGC 08-007r1.

Open Geospatial Consortium (OGC) (2005). Web 3D Service. Discussion paper. Ref no. OGC 05-019.

OpenSceneryX project (2009). <http://www.opensceneryx.com/> 10.12.2009.

OpenStreetMap (OSM) (2009). Open street map Germany. <http://www.openstreetmap.de/presse/2008-10-24-hamburg-stat.html>. 10.12.2009.

OpenStreetMap Wiki (2009). OSM proposed building attributes. <http://wiki.openstreetmap.org/wiki/Proposed_features/Building_attributes>. 10.12.2009.

OSM3D. S. Ziegler. OSM 3D Viewer (2009). <http://www.osm3d.org>. 10.12.2009.

Rossignac, J. & Borrel, P. (1993). Multi-resolution 3D approximations for rendering complex scenes. In: Falcidieno, B., Kunii, T. (Eds.), Modeling in computer graphics: Methods and applications, pp. 455–465.

Schilling, A., Basanow, J., Zipf, A. (2007). Vector based mapping of polygons on irregular terrain meshes for Web 3D map services. In: 3rd international conference on web information systems and technologies (WEBIST), March 2007, Barcelona, Spain.

Schilling, A., Lanig, S., Neis, P., Zipf, A. (2008). Integrating terrain surface and street network for 3D routing, 3D Geoinfo 08. In: 3rd international workshop on 3D geo-information, Seoul, South Korea.

Schilling, A., Over, M., Neubauer, S., Neis, P., Walenciak, Zipf, A. (2009). Interoperable location based services for 3D cities on the web using user generated content from OpenStreetMap. UDMS 2009. In: 27th urban data management symposium, Ljubljana, Slovenia.

Schneider, M., Guthe, M., & Klein, R. (2005). Real-time rendering of complex vector data on 3D terrain models. In: Proceedings of the 11th international conference on virtual systems and multimedia –VSMM, October 3–7, 2005, Ghent, Belgium, pp. 573–582.

Städtetag Nord Rhein-Westfalen (2009). 3D Stadtmodelle. Eine Orientierungshilfe für die Städte in NRW. <http://www.bocholt.de/intabox/medienarchive/fb31/3d-stadtmodelle_staedtetag_nrw.pdf>. 10.12.2009.

Statistisches Bundesamt (2009). <https://www-genesis.destatis.de>. 10.12.2009.

Wartell, Z., Kang, E., Wasilewski, T., Ribarsky, W. & Faust, N. (2003). Rendering vector data over global, multi-resolution 3D terrain. In: Proceedings of joint EUROGRAPHICS – IEEE TCV symposium on visualization, May 26–28, 2003, Grenoble, France, pp. 213–222.

Zielstra, D. & Zipf, A. (2009). Datenqualität von OpenStreetMap – Erste Ergebnisse empirischer Untersuchungen. AGIT 2009. In: Symposium für Angewandte Informatik, Salzburg, Austria.

Zipf, A., Basanow, J., Neis, P., Neubauer, S., Schilling, A. (2007). Towards 3D spatial data infrastructures (3D-SDI) based on open standards – experiences, results and future issues. In: 3D GeoInfo07, ISPRS WG IV/8 international workshop on 3D geo-information: Requirements, acquisition, modelling, analysis, visualisation, Delft, Netherlands.