



Quelle: pro motion pic / fotolia.com

GEBÄUDEDATEN VON OPENSTREETMAP – EINE ALTERNATIVE ZU AMTLICHEN GEOBASISDATEN?

Von Carola Kunze

Stadtentwicklungsanalysen, Katastrophenmanagement, Marketing und Stadtmodelle – in allen diesen Bereichen werden Gebäudedaten verwendet. Diese sollen möglichst exakt, aktuell und kostengünstig sein. Kann das Crowdsourcing-Projekt OpenStreetMap (OSM) heute schon Daten liefern, die diesen Anforderungen gerecht werden?

Acht Jahre nach der Gründung des OpenStreetMap (OSM) Projektes vollzieht sich ein Wandel. Begonnen haben die freiwilligen Mapper mit der Digitalisierung von Straßenzügen und markanten Punkten in der Umgebung (Points of Interest). Allmählich kamen zu einer ansprechenderen Kartendarstellung flächenhafte Objekte wie zum Beispiel Wald-, Wasser- und Gebäudeflächen hinzu.

Mittlerweile haben die Gebäudegrundrisse, mit rund 61 Millionen Objekten, die rund 52 Millionen Straßenobjekte sogar zahlenmäßig überholt.

GEBÄUDEDATEN FÜR ALLE

Für die Orientierung auf einer Karte dienen dem Nutzer Gebäude als wichtige Anhaltspunkte und ergänzen

das Kartenbild somit nicht nur visuell. Eine Einzelhausdarstellung ist für große Maßstäbe bzw. Zoomstufen deshalb besonders wünschenswert, erfordert jedoch erheblich größeren Aufwand bei der Datenaufnahme. Hilfsmittel der Hobby-Kartographen von OpenStreetMap für die Aufnahme der Geobjekte sind vor allem GPS-Geräte (seit einiger Zeit auch Smartphones) und digitale Luft- und

Satellitenbilder. Die Trackaufzeichnung eines Gebäudegrundrisses wird jedoch durch die Abschattung des GPS-Signales und daraus resultierenden Abweichungen nahezu unmöglich. Eine professionelle Einmessung der Gebäude übersteigt vielfach das Know-how und die technischen Möglichkeiten der Freizeit-Mapper. Was bleibt, ist die Digitalisierung der Gebäudepolygone aus Luftbildern. An erster Stelle können dafür heute die Bing-Luftbilder genutzt werden, die Microsoft dem Projekt zur Verfügung gestellt hat. Sie liegen für einige Städte sogar mit der Aktualität vom vergangenen Jahr (2011) vor. Zusätzlich zu den Gebäudepolygonen sieht die OSM-Datenbank zahlreiche Attribute vor, wie etwa die Art des Gebäudes, seine Nutzung oder auch die Höhe und die Dachform. Über verschiedene Tools oder Schnittstellen können die gewünschten Datensätze aus der OSM-Datenbank bezogen werden. Neben proprietären Geoinformationssystemen stehen auch zahlreiche OpenSource-Programme

(z. B. QGIS, Osmosis) für die Weiterverarbeitung der so extrahierten Geodaten zur Verfügung.

OSM VS. AMTLICHE GEBÄUDEUMRISSE

Außer für private und kommerzielle Anwender spielen OSM-Daten, im speziellen die Gebäudedaten, auch im Bereich der Forschung und Entwicklung eine wichtige Rolle. Für die Analyse des Gebäudebestandes einer Stadt müssen die Gebäudedaten in ihrer Genauigkeit allerdings einem hohen Anspruch genügen. Benötigt werden flächendeckende, aktuelle und positionsgenaue Daten beispielsweise für stadtplanerische Aufgaben. Die Daten sollen es erlauben, Trends zu erkennen und sie helfen, Entscheidungen zu treffen. Bereitgestellt werden solche Daten gewöhnlich von den Landesvermessungsämtern. Sie werden dementsprechend im ATKIS (Amtliches Topographisch-Kartographisches Informationssystem) und

ALKIS (Amtliches Liegenschaftskatasterinformationssystem) bzw. in der ALK (Automatisiertes Liegenschaftskataster) geführt. Durch Einmessung und/oder Digitalisierung aus Ortho-Luftbildern erzeugen die amtlichen Vermesser hochdetaillierte Datensätze, die relativ aktuell bereitgestellt werden. Überdies dokumentieren Metadaten den in den Daten dargestellten Sachstand. Eine großräumige Nutzung dieser Daten ist jedoch noch immer mit hohen Kosten verbunden. Aufgrund der steigenden Popularität und wachsenden Datenmenge des Crowdsourcing-Projektes OpenStreetMap rücken dessen Daten als kostenlose Alternative daher immer mehr in den Vordergrund. Aber wie ist es mit Blick auf die Gebäudedaten qualitativ und quantitativ um OpenStreetMap bestellt? Einen ersten Eindruck wie genau beide Gebäudedaten übereinstimmen, liefert Abbildung 1. Sie zeigt einen Ausschnitt aus den Gebäudepolygon-Datensätzen der Innenstadt Essens, hinterlegt mit Bing Aerial Imagery.



Abbildung (1) Visueller Vergleich: Gebäudepolygon-Datensätze von OSM und aus amtlichen Quellen von der Innenstadt Essens, hinterlegt mit Bing Aerial Imagery.

Nun gilt es festzustellen, ob die nutzergenerierten OSM-Daten den Anforderungen der erwähnten Aufgabenstellungen gerecht werden. Die wichtigsten Qualitätsmerkmale sind Vollständigkeit und Positionsgenauigkeit. In einer Arbeit an der Technischen Universität Dresden, in Zusammenarbeit mit dem Leibniz-Institut für ökologische Raumentwicklung in Dresden, wurde eine Vollständigkeitsanalyse des OSM-Gebäudedatenbestandes anhand zweier Untersuchungsgebiete – Sachsen und Nordrhein-Westfalen durchgeführt.



Abbildung 2: Die stark unterschiedliche Modellierung der Gebäudepolygone sorgt für zahlen- und flächenmäßige Unterschiede in den Analyseergebnissen.

VOLLSTÄNDIGKEIT DER OSM-DATEN

Für die Bewertung der Qualität bzw. Vollständigkeit eines Datensatzes bedarf es einer Referenz. Dieser Referenzdatensatz wird für die Analyse als vollständig angenommen. Ihm steht der Vergleichsdatensatz – in diesem Fall von OSM – gegenüber. Mithilfe geeigneter Methoden findet eine Gegenüberstellung beider Datensätze statt. Bei der Gebäudedatenuntersuchung stehen Gebäudegrundrisse als Referenz für Nordrhein-Westfalen in Form von so genannten Hausumringen (HU) zur Verfügung, die aus der ALK abgeleitet wurden. In Sachsen

kommen ATKIS-Gebäudedaten zum Einsatz. Beide Bundesländer werden anhand ausgewählter Städte (jeweils eine Groß-, Mittel- und Kleinstadt) sowie einer ländlichen Region bewertet. Dabei spielen vier methodische Ansätze eine Rolle.

Um einen ersten Eindruck von der Vollständigkeit der Datensätze zu erhalten, lohnt sich ein Blick auf die statistischen Werte zu Anzahl und Fläche der Gebäudepolygone. Betrachtet man beide Bundesländer nach

diesen Kriterien, beträgt der Anteil von OSM gemessen an der Referenz, für die Anzahl übereinstimmender Gebäude ernüchternde zwölf Prozent in NRW und 14 Prozent in Sachsen. Beim entsprechenden Flächenanteil ist es hingegen nahezu das Doppelte, rund 33 Prozent in NRW und 30 Prozent in Sachsen.

Wie kommen die unterschiedlichen Werte zwischen Anzahl und Flächengröße der Gebäude zustande? Grund hierfür ist die stark un-

▶ UNTERSUCHUNGSGEBIETE	▶ ANZAHL	▶ FLÄCHE [%]	▶ CENTROID [%]	▶ ÜBERDECKUNG[%]
SACHSEN	14,5	30,7	15,3	14,4
▷ Leipzig (Großstadt)	25,2	58,8	28,9	28,4
▷ Chemnitz (Mittelstadt)	24,3	60,4	32,0	31,4
▷ Bautzen (Kleinstadt)	37,4	80,8	47,8	44,5
▷ Vogtlandkreis (Ländl. R.)	6,5	15,6	4,9	4,5
NORDRHEIN-WESTFALEN	12,6	33,3	16,4	10,1
▷ Essen (Großstadt)	30,9	84,1	53,5	52,5
▷ Münster (Mittelstadt)	5,5	28,1	8,8	8,7
▷ Lemgo (Kleinstadt)	1,2	13,0	1,8	1,7
▷ Kreis Coesfeld (Ländl. R.)	9,4	21,4	10,5	9,6

terschiedliche Modellierung der Gebäudepolygone. Gibt es in der amtlichen Vermessung klare Regeln, nach denen eingemessen, digitalisiert und attribuiert wird, folgen die OSM-Mapper keiner einheitlichen Richtlinie. Reihenhäuser werden in den Referenzdaten als einzelne Polygone wiedergegeben, in den OSM-Daten nur sporadisch (siehe dazu Abbildung 2). Dieser Unterschied wirkt sich vor allem auf die Anzahl der Gebäude aus. Ein unterschiedlicher Detailierungsgrad bewirkt, dass bevorzugt verwinkelte Gebäude oder solche mit Innenhof in OSM als viel größere Fläche auftreten als in den amtlichen Daten.

Um diese Modellierungsunterschiede auszugleichen, erfordert es objektbezogene Methoden, die die einzelnen Gebäudepolygone direkt miteinander vergleichen. Das Ergebnis ist der Gebäudeanteil, welcher bereits durch OSM repräsentiert wird. Dazu stehen zwei Ansätze zur Verfügung. Entweder testet man mit der Centroidmethode, ob der Centroid des Referenzpolygons im Polygon des Vergleichsgebäudes liegt. Oder es wird mittels der Überdeckungsgradmethode aus der Verschneidung der Polygone beider Datensätze der Überdeckungsgrad zum Referenzgebäude berechnet. Liegt dieser bei über 50 Prozent, wird das entsprechende amtliche Gebäude als durch OSM repräsentiert angenommen. Aus diesen Bilanzierungen lässt sich der übereinstimmende Anteil von OSM zur Referenz für die jeweiligen Testgebiete berechnen und miteinander vergleichen. Die Tabelle (1) stellt die Ergebnisse anhand der einzelnen Methoden gegenüber.

Unterm Strich kann der aktuelle Stand der OSM-Gebäudedaten hinsichtlich der Vollständigkeit gut bewertet werden. Dieser beläuft sich mit den Daten vom November 2011 auf einen Anteil von weniger als 50 Prozent im Vergleich zu den amtlichen Gebäudedaten. Die Ergebnisse der verschiedenen methodischen Ansätze unterscheiden sich dabei nur

wenig voneinander. Am interessantesten ist der Kontrast zwischen den ländlichen Regionen und den Städten, wenn es um die Vollständigkeit des OSM-Gebäudebestandes geht. Die untersuchten ländlichen Regionen Kreis Coesfeld und der Vogtlandkreis mit ihren Unterzentren weisen noch ein deutliches Defizit im Bereich der Gebäude auf. Im Verhältnis zur eher geringen Bevölkerungszahl gibt es dem Anschein nach nur wenige aktive OSM-Mapper. Dass dies jedoch nicht die Regel ist, zeigt das Beispiel der sächsischen Kleinstadt Bautzen, welche die zwei höchsten Werte im Vergleich der untersuchten Städte aufweist. Ausschlaggebend ist demnach oft eine aktive und engagierte Community. In ihrer Gesamtheit stellen sich die OSM-Gebäude im Bezug auf geometrische Genauigkeit, Detaillierungsgrad und Attributierung als flächenmäßig stark inhomogene Daten heraus. Dies macht die Nutzung als

alleinige Basis für räumliche Aufgaben, zum Beispiel in der Planung, bisher noch nahezu unmöglich.

VISUALISIERUNG

Interessant ist die flächenmäßige Betrachtung der Vollständigkeit in den verschiedenen Städten und Regionen. Dazu kann ein regelmäßiges Raster verwendet werden, für dessen Rasterzellen die jeweiligen Anteile berechnet und mithilfe einer Farbskala codiert werden. So ist es möglich, Gebiete besonders guter oder auch schlechter OSM-Abdeckung schnell zu identifizieren. Die Stadtzentren sind dabei die Gebiete mit besonders guter OSM-Vollständigkeit. Weiter zu den Randbezirken der Städte nimmt der OSM-Anteil deutlich ab. Das in Abbildung 3 dargestellte Beispiel zeigt über ein farbcodiertes Hexagonraster für die Stadt Münster die prozentuale Differenz zwischen OSM und

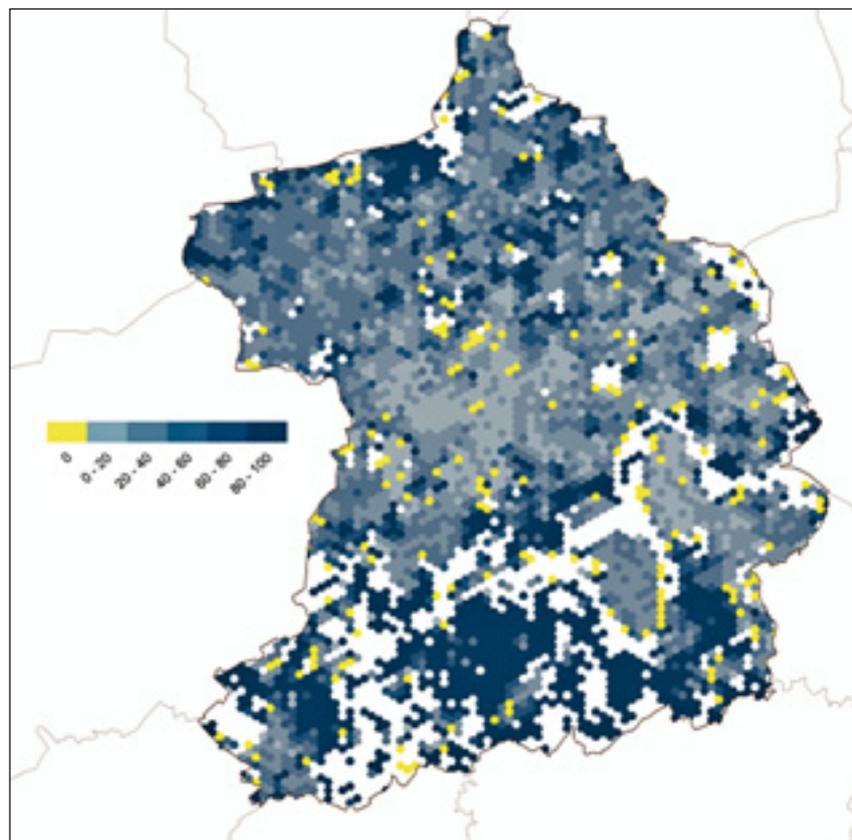


Abbildung 3: Differenz zwischen den Anzahlen der HU-Gebäude und der selektierten HU-Gebäude nach der Überdeckungsgradmethode pro Rasterzelle (in Prozent) in Münster, wobei eine Rasterzelle einen Durchmesser von 250 Metern aufweist.

den amtlichen Gebäudedaten nach der Überdeckungsgradmethode. Dieser visuelle Vergleich sowie die statistischen Ergebnisse zu den einzelnen Städten und Regionen geben auch Aufschluss über die unterschiedliche Vollständigkeit der OSM-Daten in städtischen und ländlichen Gebieten.

GEBÄUDEDATEN DOCH NICHT FÜR ALLE

Ein solch ernüchterndes Ergebnis stellt jedoch nicht das Aus für eine wissenschaftliche oder kommerzielle Nutzung dieser Gebäudedaten dar. Ein Crowdsourcing-Projekt unterliegt ständigem Wandel und Wachstum. Die Zahl der Gebäudepolygone steigt rasant und

mit jeder Bearbeitung nimmt auch deren Qualität und Fülle an zusätzlichen semantischen Informationen zu. Allein im Bundesland Sachsen sind innerhalb eines halben Jahres nach der ersten Vollständigkeitsanalyse knapp 100.000 Gebäude hinzugekommen, was einem Anteil von 35 Prozent an der Gebäudezahl vom November 2011 entspricht. Hinzu kommt, dass auch die Geobasisdaten der Vermessungsämter nicht homogen sind. Die Zusammenführung der Daten aus den aktuellen ALK-Beständen ins so genannte AAA-Modell ist in vollem Gange. Unterschiede in den Daten der einzelnen Bundesländer machen eine bundesweite Bearbeitung teilweise schwer. Das war auch der Grund für die Verwendung zweier verschiedener Referenzdatensätze. Lagen für Nordrhein-Westfalen sehr gute und aktuelle Hausumringe vor, konnten für Sachsen aus Gründen mangelnder Vollständigkeit nur die ATKIS-Gebäude verwendet werden. Das zeigt, dass auch innerhalb dieses Datensatzes trotz festgelegter Richtlinien unterschiedliche Modellierungen auftreten.

VON OSM LERNEN

Die Inhalte aus diesem freiwilligen Projekt sind inzwischen zu stattlicher Größe herangewachsen und können auf vielfache Weise nützlich sein. Dabei geht es nicht darum, die bereits vorhandenen amtlichen oder kommerziellen Datenquellen zu ersetzen. Das wird OSM vermutlich nie gelingen. Akzeptiert man diese neue Datenquelle im Bereich der Geodaten jedoch als zusätzliches Werkzeug, können daraus zahlreiche Anwendungen entstehen.

Betrachtet man die Datenaufnahme und die Möglichkeit, einfach und schnell an die eigenen Wünsche angepasste Karten zu erstellen, erschließen sich neue Bereiche. In Katastrophensituationen, wie dem Erdbeben in Haiti (2010), wurden innerhalb kurzer Zeit ganze Landstriche kartiert und die Kartendaten vor Ort von Einsatzkräften genutzt. Auch in langfristigen Projekten in Regionen, welche wenig oder schlecht

kartiert sind, kann mit einfachen Hilfsmitteln dazu beigetragen werden, die Infrastruktur etwa für Hilfeleistungen zu verbessern. Dies funktioniert beispielsweise im Projekt MapKibera in Nairobi, Kenia. OSM erhält seinen Wert so gesehen weniger aus seiner Eigenschaft als Datenbank als vielmehr aus seiner Fähigkeit ein weltweit verfügbares und schnelle erlernbares Werkzeug für beliebige Kartierungen anzubieten.

Gleichzeitig kann OSM auch als gutes Beispiel für weitere Projekte mit nutzergeneriertem Inhalt dienen. Neben dem Datenschema ist auch die bestehende Infrastruktur des OSM-Projektes für Anwender im Geoinformationsbereich interessant. Auch viele der Verarbeitungswerkzeuge liegen als Open-Source-Lösungen vor und können dadurch bestens an die eigenen Aufgabenstellungen angepasst werden.

Ist also der Gebäudebestand von OSM zum jetzigen Zeitpunkt noch nicht als alleinige Quelle für komplexe Aufgaben ausreichend, besteht trotzdem die Möglichkeit, diese Daten ergänzend einzusetzen. In den eigenen Datenbeständen können durch einen Vergleich beider Datensätze eventuelle Lücken identifiziert werden. Auch wenn die Aufnahmeaktualität in OSM stark variiert, reagieren viele Mapper sehr schnell auf neu hinzugekommene Gebäude und halten die OSM-Daten auf einem sehr aktuellen Niveau. Ob die Vervollständigung der Gebäudedaten ähnlich schnell wie bei den Straßendaten vonstatten geht, wird sich zeigen. Fest steht, dass mit jedem neuen Anwender auch die Bedeutung der freien Geodaten von OpenStreetMap steigt. ◀

GEBÄUDEDATEN AUS OSM EXTRAHIEREN:

- ▷ **Download:** Bezug der OSM-Daten über Drittanbieter wie zum Beispiel die Geofabrik, über vorhandene Plug-ins und Schnittstellen in GIS-Lösungen oder direkt über die OSM-API (API v0.6)
- ▷ **Import:** Je nach Download-Variante Import in ein GIS oder ein Datenbanksystem
 - Für ArcGIS bietet ESRI das Add-on „ArcGIS Editor for OpenStreetMap“ an (Import der OSM-Datei in eine FeatureClass).
 - Über die Anwendung Osmosis können OSM-XML-Dateien (.osm) in eine PostgreSQL-Datenbank mit der PostGIS-Erweiterung abgelegt werden.
- ▷ **Selektion:** Extraktion aller Gebäude aus den Polygon-Layern über das Tag „building = “. Suche aller Objekte, die einen Wert für den Schlüssel building haben z.B. yes, school, etc.
- ▷ **Weiterverarbeitung:** Nutzung der Geoverarbeitungswerkzeuge für weitere Analysen

AUTORIN UND KONTAKT:

Carola Kunze
TU Dresden
Institut für Kartographie
carola.kunze@googlemail.com

Weiterführende Links:

- 🔗 www.openstreetmap.org
(die Karte)
- 🔗 wiki.openstreetmap.org
(das Wiki mit allen wichtigen Hinweisen zu OSM)
- 🔗 download.geofabrik.de
- 🔗 wiki.openstreetmap.org/wiki/Osmosis
- 🔗 www.postgresql.org
- 🔗 postgis.refractory.net