IDENTIFYING BUILDING TYPES AND BUILDING CLUSTERS USING 3D-LASER SCANNING AND GIS-DATA

Hauke Neidhart, Monika Sester

Institute of Cartography and Geoinformatics (IKG), University of Hannover,
hauke.neidhart@ikg.uni-hannover.de

KEY WORDS: Laser scanning, DEM/DTM, Building, Extraction, GIS, Classification, Combination

ABSTRACT:

Power authorities are highly interested in figures that indicate the energy requirements and especially the heat requirements on a local, regional and country-wide level. Such numbers are needed for their planning of new sites of power plants or for planning alternative energy modes.

Existing methods for estimating those requirements heavily rely on local sampling methods as well as on the use of statistical estimates and models. The traditional way of acquiring area wide data is to use statistics and punctually acquired data and extrapolate it to wider areas. E.g. several districts of a city are investigated based on aerial photos and classified into different building and settlement typologies; the cities, in turn are classified according to certain types, which in the end will lead to a country wide statistics.

In order to determine more accurate base information, in this project we are using laser scanning as a basic data acquisition method to determine building volumes, i.e. the volumes to be heated. This is due to the fact that laser scanning potentially allows for an area-wide data capture, and also has a high potential of automated data analysis and interpretation. The heat demand of an individual building depends primarily on its age and its type. Therefore, in order to assign head demands to individual buildings measured from laser scanning, the building type first has to be inferred from the available geometric characteristics.

The paper will present the results of the automatic extraction of building volumes, and concentrates on the identification of the given building and settlement types that can be used to link the building volumes with specific heat coefficients. The results achieved with our approach will be compared with results derived in the traditional way.

1. INTRODUCTION AND OVERVIEW

Our work is part of a project on pluralistic heat supply ("Pluralistische Wärmeversorgung") which is funded by the AGFW (Arbeitsgemeinschaft Fernwärme, e.V.). AGFW is an organization of energy and service providers which are engaged in local and district heating.

One of the goals of this project is to detect locations where local and district heating can compete with traditional heating by electricity, gas or oil. For that purpose, model calculations are performed which in turn need highly detailed information on the heating demand. However, existing information is often out of date or not available on an area-wide basis.

Thus, we aim to derive this information from different data sources. Since the heating demand of buildings is correlated with the building volume, our first goal was to extract building volumes (Neidhart & Brenner, 2003). In a second step, these can be combined with additional information such as specific heat coefficients which depend on the building type and year of construction.

There has been a huge amount of research in the field of automatic extraction and reconstruction of man-made objects, including buildings, see e.g. (Baltsavias et al. 2001). For example, Weidner (1997) uses laser scan data to extract buildings. Using a segmentation of a normalized digital surface model (DSM) the locations of buildings are detected. From this, the ground plans are reconstructed. Brenner (2000) describes the reconstruction of 3D-buildings from laser scan data and ground plans, leading to detailed roof topologies. However, in our case a detailed reconstruction of the building's geometry seems not to be necessary, as we are interested mainly in the volume and not in the exact shape. Other geometrical features such as roof area and slope might be interesting at a later stage, for example to derive assumptions about the year of construction.

DSMs from laser scanning are well suited to derive building volumes as they generally preserve jump edges quite well and are easier to use in automated methods as compared e.g. to aerial images. There are many different algorithms to derive a DTM from a given DSM. For example, Masaharu and Ohtsubo (2002) divide the area into small tiles and select the lowest points. In a further step, it is verified if these represent the terrain. Then the initial DTM is created. At the moment this method can only be used in flat terrain. Briese et al. (2002) use robust methods to classify the original points into terrain and off-terrain points.

In order to obtain a precise definition of the terrain, additional data sources can be integrated. For our studies, we explored the use of ATKIS and ALK datasets. ATKIS is the Authoritative Topographic Cartographic Information System in Germany (ATKIS, 2003). It contains information on settlements, roads, railways, vegetation, waterways, and more. However, for our studies we only used the roads layer. ALK is the digital cadastral map containing information on parcels and buildings (AdV, 2003). Again, we used only a small part of the available information, namely the ground plans of buildings.

After the extraction of the building volumes from the laser scanning data in the first step, the buildings have to be related to heat demand. Heat requirement not only depends on the building volume alone, but also on other properties, especially the type of the building. Furthermore, also the outer surface of the building, where the heat can emit, plays an important role. Depending on the age of the building assumptions on the insulation of the outer surface can be made and introduced in the estimation of the building-specific heat requirement.

There are different typologies for residential buildings, for which specific heat coefficients are known. These typologies classify the buildings by means of size (one-family house, multi-family residence) and age. This typology, however, describes the characteristics of the buildings on a level which is made for human interpretation. Therefore, the main task is to set up rules that can infer the building type from a set of observable and measurable building characteristics, and thus link the mere geometric data with specific heat coefficients. Besides geometric information of the building itself, namely ground plan and volume, also their relative arrangement among each other and to roads have to be taken into account.

The idea in our approach is to extract the geometric characteristics using all available data (cadastral map, topographic data, streets, …). The characteristics relate to the building itself, i.e. length, width, width-to-length-ratio, roof type, etc. and they regard also the context, i.e. distance to neighbouring buildings and roads. We use the given building typologies and determine significant characteristics for every type. This will be achieved using a Machine Learning approach, that automatically derives the discriminating and characterising attributes of a given classified data set. Furthermore, we will also identify settlement areas with similar characteristics using a clustering approach. Based on such a settlement typology, the estimation of the heat demand for very large areas will be eased.

The structure of the paper is as follows: after a description of the workflow, the building and settlement typology is introduced. Then methods for identifying building types and settlement types will be presented. Finally, results are shown that can be achieved using this method. A summary and outlook on future work concludes the paper.

2. WORKFLOW

The way from the raw data to the heat demand consists of the main steps.

1. determination of building volumes
2. classification into different building types and
3. calculation of the heat demand using the volume and the specific heat coefficient.

The first step is the determination of building volumes. For the determination of building volumes the main data are the laser scanning data. But we also have tried out different combinations of laser scanning and other GIS data. For the following steps the combination of laser scanning and the cadastral map is useful. With the attributes from the cadastral map it is possible to link additional data with the building volumes.

The next step is to classify the building volumes to different building types. Each building type has a specific heat coefficient. A lot of preparatory work has to be done because there is no direct link between the geometry of the building volume and the building types. To do so, the buildings are classified into different types according to a given typology, using only geometric properties available.

In the last step the volumes are combined with the specific heat coefficient from the building typology. Then the heat demand can be calculated and represented in a map.

For test purposes, the results achieved with the method described here are compared with the data from a heat atlas, that was available and was acquired with conventional methods. A link between our results and the data from the heat atlas is possible using the addresses of the buildings which are present in both data sets.

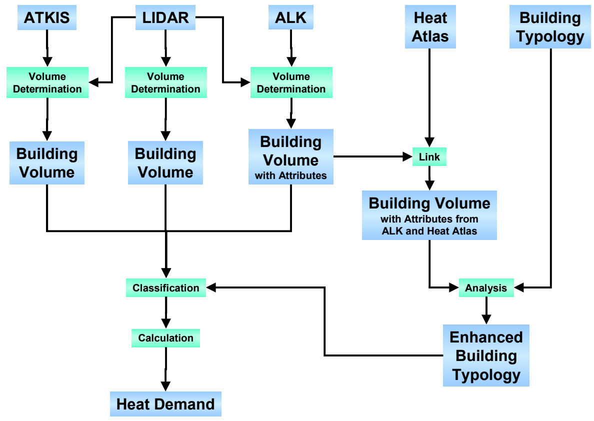

Figure 1: Workflow to derive heat demand map from original spatial data.

3. BUILDING AND SETTLEMENT TYPOLOGY

A building typology classifies residential buildings according to their size and the age. The used building typology is shown in Table 1. Depending on the size and the age each type has a specific heat demand.

From the one-family house towards the large more-family house the buildings become more compact. Therefore the specific heat demand is reduced.

The limits for the building ages relate to different building regulations which concern the quality of heat insulation. The newer regulations demand better insulation. So newer buildings have lower specific heat demands.

	specific heat demand [kWh/m²*a]			
building age	**one-family house**	**row-house**	**small more family house**	**large more family house**
until 1918	205,5	199,6	187,8	124,4
1919-1948	206,0	173,2	151,1	168,9
1949-1957	252,4	162,5	174,9	140,6
1958-1968	185,3	161,8	179,7	160,4
1969-1977	155,4	146,2	136,6	139,4
1978-1984	139,5	133,3	109,0	105,9
1985-1995	139,5	115,5	81,4	75,8
1996-2000	105,9	106,2	95,4	86,4

Table 1: Building Typology.

One can easily observe that the description of the building types only contains information about the heat demand and no geometry. For statistical methods this is sufficient because cities often have statistics about the types of buildings in their town. However, if this is not available, the evaluation has to be done manually using aerial photographs. In any case, with this statistical methods there is no information about the spatial distribution of the heat demand.

One thing one should remember is, that this building typology only considers residential buildings. So it can only used in residential areas and not for instance in industrial areas.

A settlement typology describes the combination of different building types within a certain area, e.g. a city centre consist of blocks of great more family houses and a suburb consist of one-family houses or row-houses.
Some settlement types are more suitable for local and district heating. For instance, in dense areas it is very expensive to lay the pipelines, so these areas are not suitable for this kind of heating.

4. IDENTIFYING BUILDING TYPES

From the laser scanning data we only obtain geometric information about the building.
The main task is to create a link between the geometry and the building types. This is a classification task, that relates object attributes with a certain object class. To this end, different classification methods can be applied. Using supervised classification, one starts with training data, which contain classified examples. In our case, in a first step the building volumes are combined with a heat atlas which contains relevant data to identify building types, which were used as training data.

We have to decide which attributes should be used. The most characteristic attributes are the height, length, width and area.
An analysis delivers an enhanced building typology which includes geometric values. This enhanced building typology then can be applied to other regions

Below different methods to determine the building type are described.

4.1 Approximate values

In this method we start with approximate values. These values can be obtained from experience. From statistical data we can obtain some information about certain building types, e.g. number of floors, number of apartments and the size of apartments. Besides the length and width become greater from one-family house to the high tower building.

building type	area [m²]	height [m]	length [m]	width [m]
one-family house	90-115	3-7.5	13-15	7-8.5
row-house	70-110	3-7.5	10-12	7-8
small more-family house	90-150	6-11	10-17	10-11
large more-family house	140-260	10-18	14-24	10-16
tower block	400-900	28-45	25-65	20-30

Table 2: Initial values for building type classification.

Another way to determine the values is to manually select buildings which types are known. The Table 2 shows the initial values.

With these values from the training data set all buildings of the whole data set are classified. After the classification a visual inspection takes place. Then, the values are iteratively improved until all buildings are correctly classified.

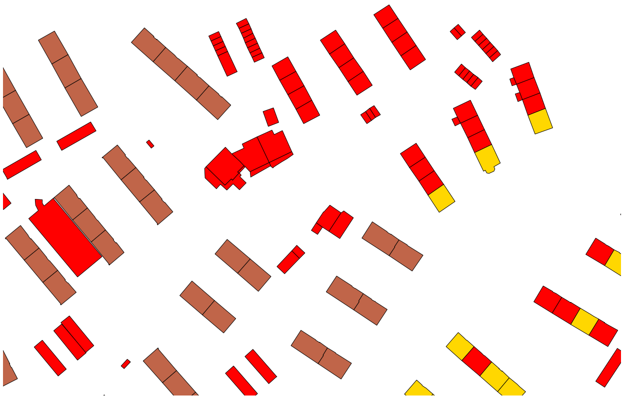

Figure 2: Result of the classification with initial values. Brown buildings are classified as large more-family houses. Yellow buildings are classified as small more-family houses. Red buildings are not classified.

Figure 2 shows the first classification with the approximate values. There are small and large more-family houses in this area. Many small more-family houses (red) are not assigned to the right building type (yellow). This is due to the fact that the initial values for the area and the height were a little too small. The large more-family houses are already assigned to the correct class (brown).

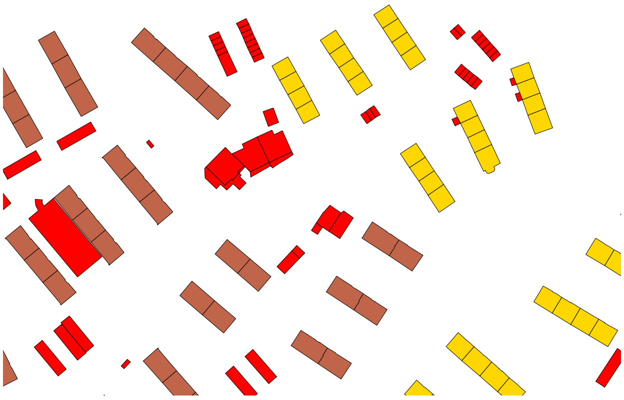

Figure 3: Result of the classification with adapted values. Brown buildings are classified as large more-family houses. Yellow buildings are classified as small more-family houses. Red buildings are not classified.

In figure 3 are the results for the adapted values from table 3. Now most buildings are correctly classified. Only garages, subterranean garages and commercial buildings are not assigned to a class.

building type	area [m²]	height [m]	length [m]	width [m]
One-family house	90-115	3-8.5	13-15	7-8.5
row-house	60-110	3-8.5	8-12	7-9
small more-family house	90-160	6-12	10-17	10-11
large more-family house	140-280	10-18	14-24	10-16
tower block	280-900	27-60	20-75	17-30

Table 3: Adapted values

This method needs a lot of manual processing and therefore is very time-consuming. Furthermore, it has to be adapted to different settlement types of a city, e.g. the values differ from city centre to suburban area.

4.2 Clustering

In the second approach we use an unsupervised classification method, namely clustering for the automatic search of groups with similar attributes. Instead of adjusting the values to classify most of the buildings, all buildings are used and every building is assigned to a cluster.

In this step we used the program package WEKA (WEKA, 2000). The table 4 shows the mean value and the standard deviation for some clusters.

After all the buildings are assigned to a cluster, each cluster has to be assigned to the appropriate building type. This is less time-consuming than the iterative method described in 4.1.

Cluster	Height [m]	Area [m²]	Length [m]	Width [m]	Building Type
	Mean (StdDev)	Mean (StdDev)	Mean (StdDev)	Mean (StdDev)	
0	3.78 (1.04)	20.42 (3.08)	5.85 (2.64)	3.74 (2.10)	Garage
1	10.58 (2.07)	253.69 (19.44)	19.94 (1.56)	15.23 (0.90)	LMFH
2	14.31 (2.40)	177.74 (15.85)	18.10 (0.92)	10.36 (0.89)	LMFH
3	9.82 (1.10)	174.50 (14.81)	17.29 (1.19)	10.30 (0.70)	LMFH
…					
12	7.33 (1.47)	125.10 (15.61)	13.19 (1.37)	10.60 (0.71)	SMFH
13	8.46 (0.40)	91.64 (4.91)	10.74 (0.39)	8.96 (0.15)	SMFH
…					
18	41.27 (4.79)	368.42 (13.70)	25.65 (1.41)	18.39 (0.91)	Tower block

Table 4: Cluster and assigned building types.

This method also delivers a more detailed building typology because different characteristics for the same building type are considered.

Figure 4 shows an example of the clustering On the left are blocks of large more-family houses. In the upper left corner are one-family houses (pale red). On the right are office buildings and stores.

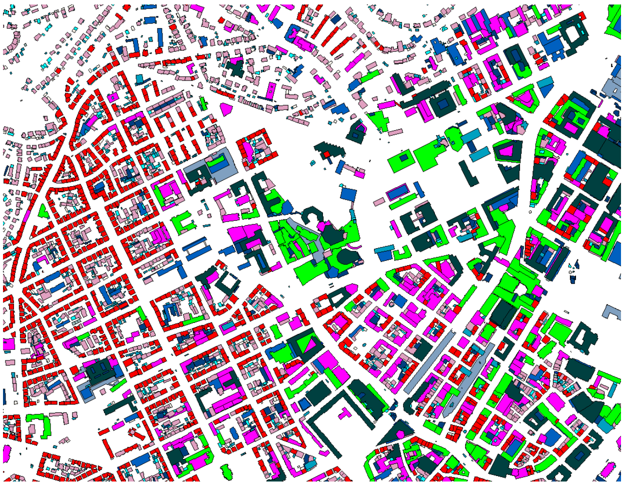

Figure 4: Results for the city centre of Stuttgart.

It also reveals that there are larger areas with the same building types.

5. IDENTIFYING SETTLEMENT TYPES

The settlement types are composed of certain building types and they vary in density and the arrangement of buildings.
To identify these different settlement types, clusters of objects with similar properties and characteristic spatial distribution have to be found. One possibility to do so is to use spatial clustering (Anders, 2002). Here, not only the similarity between the object is taken into account, but also the similarity in the spatial density. Also here, however, the final assignment to a certain settlement type has to be done manually after the automatic clustering.

The other possibility is to use information from ATKIS. The streets separate the surface into many small areas. Mostly these small areas coincide with the building blocks. Furthermore, these areas are assigned specific settlement types. Depending on this type and further block characteristics, e.g. building types, density or average distance to road, the assignment to settlement types from the settlement typology can be done. Then neighbouring blocks with similar characteristics can be merged.

6. RESULTS

Figure 5 shows the result of our method using laser scanning, ground plans from ALK and specific heat coefficients from the building typology. Figure 6 shows the heat demand from an existing heat atlas. The values differ about 10 to 20 per cent.

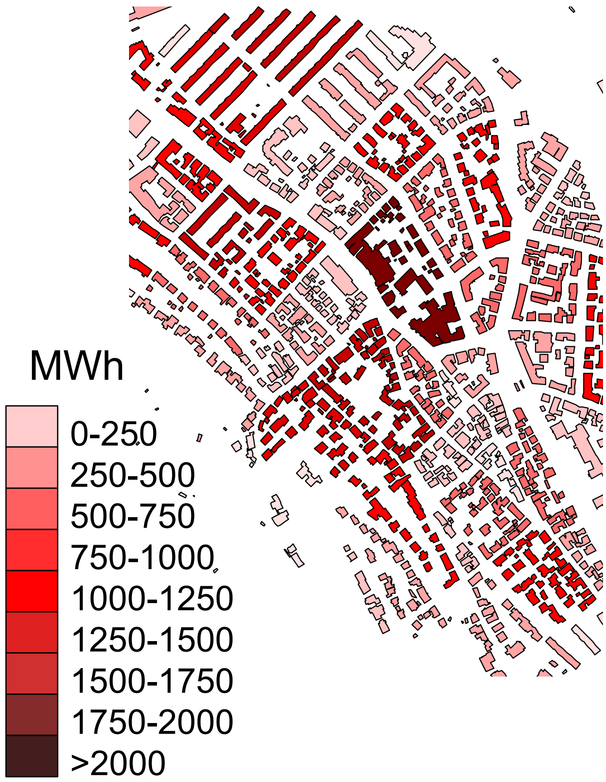

Figure 5: Heat demand determined by using building volumes and specific heat coefficients.

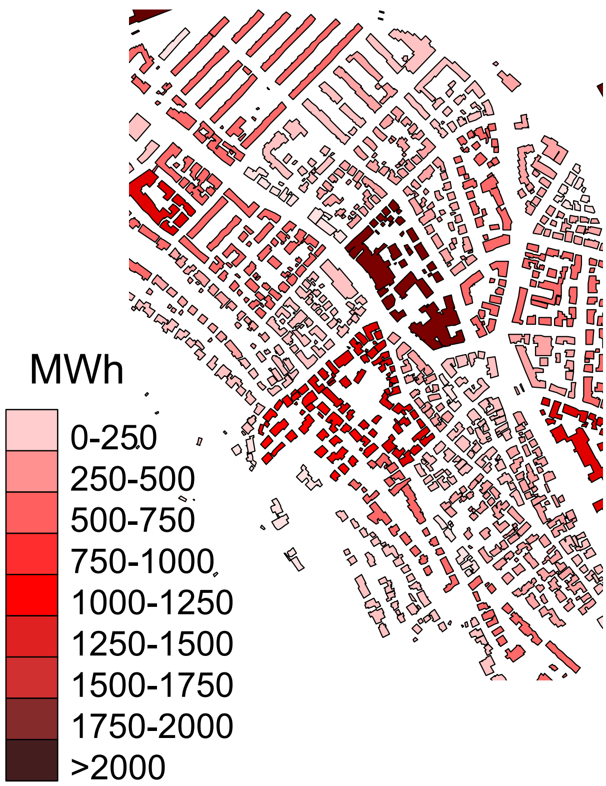

Figure 6: Heat demand from an existing heat atlas.

7. CONCLUSION AND OUTLOOK

The result shows that the method is able to deliver similar values like a heat atlas and can be used for regions that do not have a heat atlas.

Currently, we are trying to improve our method in several ways. Firstly, we are extending the set of attributes for the individual buildings, by also taking relations into account, e.g. the distance to the street, or the distance to neighbouring buildings.
At the moment the building typology only considers residential buildings. In future we will use other training datasets which also have building types for other buildings, e.g. schools or industrial buildings.

The next step is to analyze the settlement types with regard to find suitable regions for local and district heating. To this end, not only the individual settlement areas of similar type have to be analyzed, but also other factors have to be taken into account, e.g. existing pipelines, or the size of the area and distribution of the buildings, which directly is linked to the costs for installing the pipelines.

ACKNOWLEDGEMENTS

We gratefully acknowledge the funding by the AGFW (Arbeitsgemeinschaft Fernwärme, e.V.). The laser scanning data was provided by the Landesvermessungsamt Baden-Württemberg, building ground plans by the Stadtmessungsamt Stuttgart.

REFERENCES

AdV, 2003. Arbeitsgemeinschaft der Vermessungsverwaltungen der Länder der Bundesrepublik Deutschland, http://www.adv-online.de/english/products/lk.htm, accessed April 2004.

Anders, K.-H., 2002: Parameterfreies hierarchisches Graph-Clustering Verfahren zur Interpretation raumbezogener Daten, Dissertation, Hannover.

Atkis, 2003. http://www.atkis.de/, accessed April 2004.

Baltsavias, E. P., Gruen, A., van Gool, L., 2001: Automatic Extraction of Man-Made Objects from Aerial and Space Images III, Balkema Publishers.

Brenner, C., 2000: Dreidimensionale Gebäuderekonstruktion aus digitalen Oberflächenmodellen und Grundrissen. PhD Thesis. Universität Stuttgart, Institute for Photogrammetry.

Briese, Ch., Pfeifer , N., Dorninger, P., 2002: Applications of the Robust Interpolation for DTM Determination. Symposium ISPRS Comm. III, Graz, 9 - 13 September 2002. International Archives of Photogrammetry and Remote Sensing, Volume XXXIV / 3A.

Masaharu, H., Ohtsubo, K., 2002: A Filtering Method of Airborne Laser Scanner Data for Complex Terrain. Commission III, Working Group III/3.

Neidhart, H., Brenner,C., 2003: Automatic Calculation of Building Volumes for an area-wide Determination of Heat Requirements, ISPRS Commission IV Joint Workshop "Challenges in Geospatial Analysis, Integration and

Visualization II", Proceedings, September 8 - 9, 2003, Stuttgart, Germany, pp. 137-142.

Weidner, U., 1997: Gebäudeerfassung aus Digitalen Oberflächenmodellen. PhD Thesis. Rheinische Friedrich-Wilhelms-Universität zu Bonn, Institut für Photogrammetrie. Deutsche Geodätische Kommission, Reihe C, Heft Nr. 474.

WEKA, 2000: "Data Mining: Practical machine learning tools with Java implementations," by Ian H. Witten and Eibe Frank, Morgan Kaufmann, San Francisco, 2000.