

RESEARCH

Open Access



# Data completion in building information management: electrical lines from range scans and photographs

Ulrich Krispel<sup>1\*</sup> , Henrik Leander Evers<sup>2</sup>, Martin Tamke<sup>2</sup> and Torsten Ullrich<sup>1</sup>

## Abstract

**Background:** The concept of building information management (BIM) is based on its holistic nature. This idea pays off, if all relevant information is fused into one consistent data set. As a consequence, the completeness of data is vital and the research question on how to complete data automatically remains open.

**Methods:** In this article we present a data completion technique based on knowledge management. We encode expert and domain knowledge in a generative system that represents norms and standards in a machine-readable manner. The implementation of this approach be used to automatically determine a hypothesis on the location of electrical lines within indoor range scans.

**Results:** The generative paradigm can encode domain expert knowledge in a machine-readable way. In this article we demonstrate its usage to represent norms and standards.

**Conclusions:** The benefit of our method is the further completion of digital building information models – a necessary step to take full advantage of building information modeling.

**Keywords:** Data completion, Formal language, Shape grammar, Knowledge management

## Background

Studies conducted on the productivity of the construction industries show that the industry could improve efficiency using a standardized workflow across its stakeholders, particularly in the field of electrical construction companies. According to these studies the implementation of building information modeling (BIM) is a solution (Bernstein and Jones 2012; Malleson 2015); its application would reduce conflicts and improve coordination. A survey conducted in 2015 showed that the implementation of BIM has not yet reached the majority of companies (Braun et al. 2015). The study by Hanna et al. (Hanna et al. 2014) points out the deficits of actual implementations of BIM in the electrical construction field. Other sectors of the construction industry have increased their productivity using BIM (Øbro 2015). The globally accepted Industry Foundation Classes (IFC) standard (ISO 2013) includes a

complete set of electrical entities. The lack of BIM implementation in the field of building electricity is surprising, as shown by a survey conducted in the United States (Azhar 2009). This study shows a ranking of BIM features used, e.g. clash detections, visualization of electrical design or space utilization. Out of the partaking companies, 21% use BIM and report positive savings in time and cost. The remaining 79% respond that they are not using BIM due to not knowing about BIM, lack of technological experience, software incompatibility, and implementation costs. In addition, 59% of the companies which actually use BIM have only three or less years of experience with this technique. Nevertheless, these studies show that there is an interest in the utilization of BIM for electrical construction. There seems to be a hesitation of participants of the electrical construction sector to engage with BIM. This is at least partially grounded in the fact that BIM data is generally not available for projects that deal with the existing building stock which accounts for the biggest share of the building market, e.g. in the European Union for 75% (Atanasiu et al. 2011). Most of buildings date to a pre-digital area (Rapf et al. 2015). Project planning

\*Correspondence: ulrich.krispel@fraunhofer.at

<sup>1</sup>Graz University of Technology & Fraunhofer Austria Research GmbH, Graz, Austria

Full list of author information is available at the end of the article

for such buildings requires extensive and expensive documentation of the as-built state. In Europe, 61% of the overall building stock consists of pre-1980 buildings. The need to upgrade these buildings in terms of energy efficiency (European Parliament 2002) demands investments in a range between 584 and 937 billion euro until 2050.

The lack of information on the existing building stock makes planning and execution of upgrade and renovation tasks challenging and resource demanding. Even if 2D plans, 3D models, or more advanced facility management systems exist, these are rarely up to date and require adjustments to reflect the true state of the built reality. To acquire this state, terrestrial laser scanners (TLS) currently provide a means for quick three-dimensional acquisition and documentation of building stock (Tamke et al. 2016). A way to automatically document the electrical lines in existing buildings in BIM formats is unknown. Such an approach would benefit all stakeholders that deal with the planning of buildings:

- architects, engineers and interior designers can consider the existing electrical system in their design and reduce collisions of added elements and existing structures;
- electrical planning companies are provided with a good base for planning of electrical systems,
- electrical construction companies can reduce risks in their offers for building jobs,
- building owners have an overview of the existing systems and a tool to control whether planned electrical systems collide with existing elements and systems in the building.

The current situation regarding the documentation of electrical systems in existing building stock is however challenging, as it is, if at all, only present as a 2D drawing that does not provide information about the concrete position of cables in three dimensions. Figure 1 shows a typical example of such a 2D drawing as it is produced in the planning phase, and later used for work on site. It contains the necessary information for workers to perform the work in the building, such as the position of switches and fuse boxes, as well as the principal connections between them. The workers themselves know about the height at which switches are to be mounted in walls, and the way that cables shall be laid. These rules are set up by national authorities, and sometimes harmonized on a bigger scale, e.g. on the European level. Technical standards, such as the DIN 18015 in Germany (see also Fig. 2), or the “Stærkstrømsbekendtgørelsen”, the technical standard in Denmark (Vækstministeriet 2006) specify so-called “installation zones”, which should be preferred for cable routing. Workers might adapt the routing to special features on local level, but will generally follow the concept of the norms.

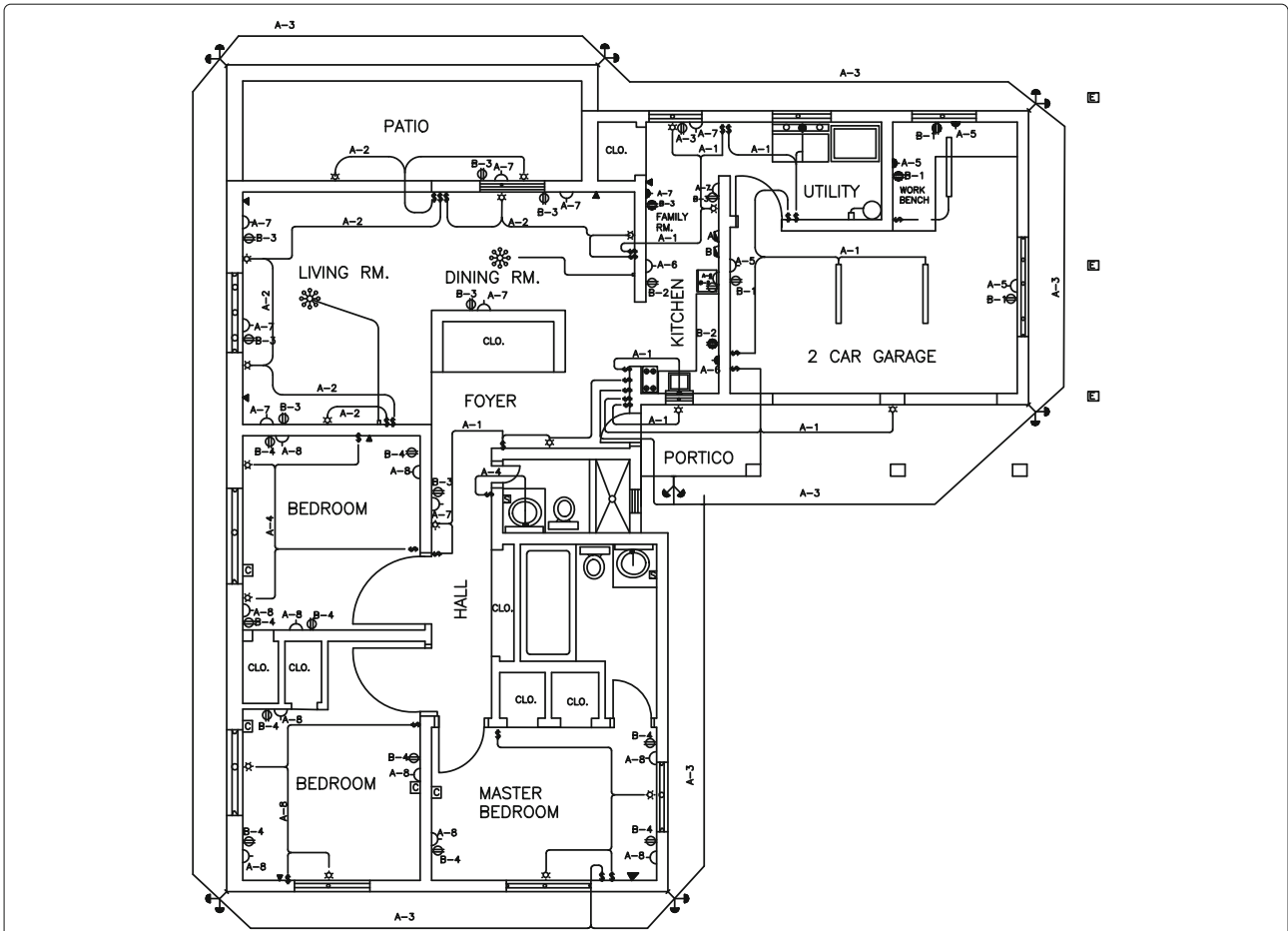
The general compliance of the electrical installation to the framework built by norms provides a relatively secure ground for professionals to estimate where wires could be routed, but assured data is usually not available. This situation provides the base scenario of this paper and other recent research: In the work of Petkova et al. (Petkova and Rüppel 2014), the location of sockets in 2D floorplans of buildings were used to estimate the amount of copper wire in the walls, a valuable resource and a real asset in the evaluation of a buildings value before demolition. This method integrates the prior knowledge of wiring placement directly into the algorithm, however, the norms differ between countries and been revised over time. In contrast, the method presented in this paper proposes to encode this knowledge using a knowledge management approach.

### Knowledge management

The main idea of knowledge management is to represent prior knowledge and expert knowledge in a self-contained, machine-readable manner, independent of the usage scenario. A commonly used approach to represent knowledge is based on machine learning. It “teaches” a machine to solve a problem using a training dataset; i.e. a set of input data with corresponding, labelled output data. Many computer vision problems, such as segmentation, detection, recognition, and matching, can be solved by machine learning techniques (Bishop 2007). Although many problems can be solved by machine learning, its application has some important limitations. Most machine learning techniques require a sufficiently large, labelled training dataset – a crucial precondition, which may be elaborate to meet. From the knowledge management point of view another problem of machine learning algorithms may arise: the “learned” knowledge can seldom be imported or exported in a human-readable way. As a consequence, prior knowledge, which may be available in form of specifications, standards and norms cannot be reused and the question what has been learned remains unanswered. Therefore, we chose a rule-based approach to represent knowledge (Schinko et al. 2010). The main advantage of procedural modeling techniques is the included expert knowledge within an object description; e.g. classification schemes used in architecture and civil engineering can be mapped to procedures. For a specific object only its type and its instantiation parameters have to be identified.

The paradigm of procedural and generative modeling is to describe an object by a sequence of operations analogue to an algorithm. The advantages of the generative approach are (Havemann et al. 2012):

- complex models become manageable through a few high-level parameters,



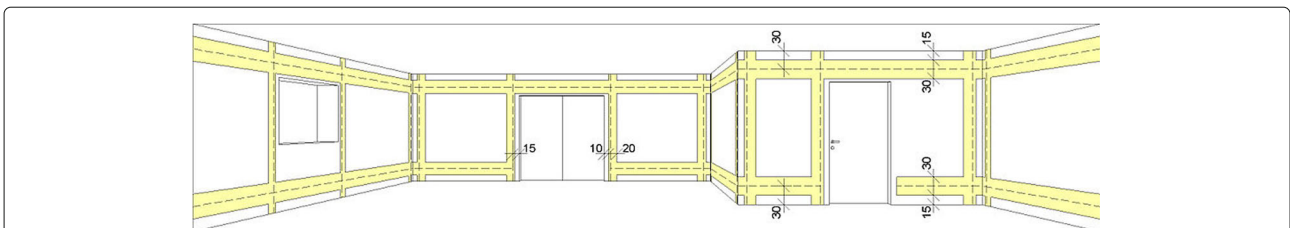
**Fig. 1** Electrical plans. This image shows a typical electrical plan for a small one family house, such plans are the common interface to communicate the locations of sockets and switches to on-site construction. This image was retrieved from (OAS/USDE)

- models are easier to store and to transmit, as only the process itself is described, not the processed data, i.e., the end result,
- changeability and re-usability of existing solutions to modeling problems can be very much improved, and
- the smaller parameter space can lead to much better results in model-based indexing and retrieval.

Figure 3 illustrates an architectural use-case (Thaller et al. 2013). The “building block” of a window in

neoclassical style has been programmed in a scripting language. Each variation of the high-level parameter set passed to the script results in a variation of the generated geometry. This example shows that all objects with well-organized structures and repetitive forms can easily be described procedurally. The procedurally encoded knowledge represents the complete family of windows; i.e. the neoclassical style of windows.

Procedurally encoded knowledge can be used to solve inverse problems such as model segmentation, semantic



**Fig. 2** Technical Specifications. Installation zones for electrical installations based on DIN18015. Zones are defined by a minimum and maximum distance from elements like floors, doors, windows, etc. Image from (Wikimedia Commons 2014)



**Fig. 3** Generative Building Blocks. Modeling of a family of similar structures can be facilitated using a generative approach, shown by this architectural example (Thaller et al. 2013): representative windows, respectively highlighted by a red and blue border, are modeled using a scripting language. Variations of these window types are generated by only replacing assets and adjusting proportions

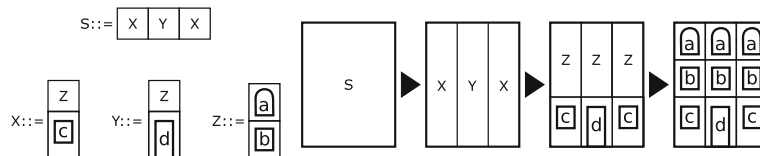
recognition and model completion. The article on “Irregular Lattices for Complex Shape Grammar Façade Parsing” (Riemenschneider et al. 2012) demonstrates a reconstruction pipeline using prior knowledge encoded in a generative representation. The authors use a shape grammar to specify the rules and regularities of façades. In a reconstruction task with incomplete data, the knowledge about possible façades, i.e. about possible outcomes of a reconstruction process respectively, reduces the search space of the reconstruction significantly. An overview of the reconstruction pipeline is illustrated in Fig. 4.

Such grammar systems were originally developed in formal language theory, and are commonly used in compiler construction. They are also a common tool used by generative modeling techniques (Thaller et al. 2013). For a review on generative modeling techniques we refer to (Krispel et al. 2015). The proposed grammar ruleset can be adapted to reflect changes in norms, or encode different norms, typically by a domain expert. The method is implemented as a pipeline that uses indoor 3D scans of buildings as a starting point. These scans consist of range scans and image information. From this data, a 3D representation of the scanned spaces is generated via a preprocessing step, and is further processed to determine the configuration of installation zones and a possible wiring hypothesis for the acquired data.

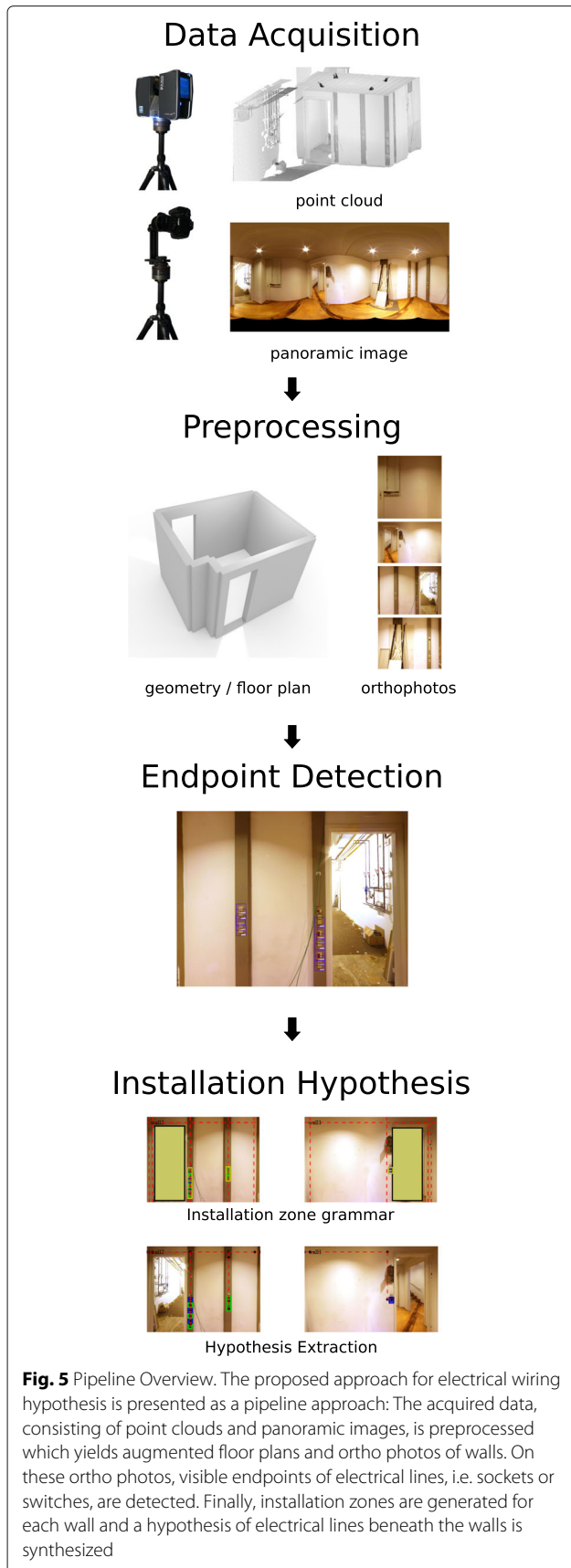
**Methods**

The proposed approach is divided into four main concepts: data acquisition, data preprocessing, the detection of visible endpoints of electrical installations (sockets and switches), and eventually the hypothesis of installation zones and a possible wiring inside the walls.

Figure 5 shows a visual overview over the whole pipeline. The first part, data acquisition, is concerned with the acquisition of measurements in the form of range scans and image information, i.e. photographs. In the preprocessing stage, a coarse room layout and geometric model is reconstructed from the point cloud. Furthermore, orthographic projections of the image information of each wall are generated. In these orthophotos, visible points of electrical lines, e.g. sockets or switches, are detected using a classification algorithm that has been pre-trained on a set of exemplar images of sockets and switches. From these detections and additional information that were extracted from the reconstructed model, i.e. the adjacency of walls and the location of openings (windows and doors), the location of preferred installation zones for electrical wirings are generated in consideration of norms and requirements. By means of the extracted wall adjacency, the installation zones inside all walls are connected to one graph that represents all possible routes along preferred zones. From this graph, a final hypothesis



**Fig. 4** Grammar based building façade reconstruction. A grammar that specifies a simple variant of buildings is specified as a twodimensional shape grammar with the rules S, X, Y and Z (left). Applying the rules yields a façade layout (right). Such a grammar can be facilitated for the inverse problem (Riemenschneider et al. 2012), in which a meta ruleset is given that specifies all feasible façades, and a concrete best matching parse tree from this feasible façade family is determined for a given façade photograph



of wall wirings is extracted under the assumption that all detected points should be connected using the least amount of resources, i.e. cable length.

Although the final result of the pipeline is non-visual (i.e. a semantically enriched BIM model), all steps produce a visual output which is used for debugging the pipeline, for the verification of the results and for publication purposes. The point clouds have been rendered using CloudCompare<sup>1</sup>; the room geometry is exported as X3D<sup>2</sup> and rendered with Blender<sup>3</sup>; the endpoint detection returns JPEG for visual inspection; the hypothesis modules generates SVG files for visualization.

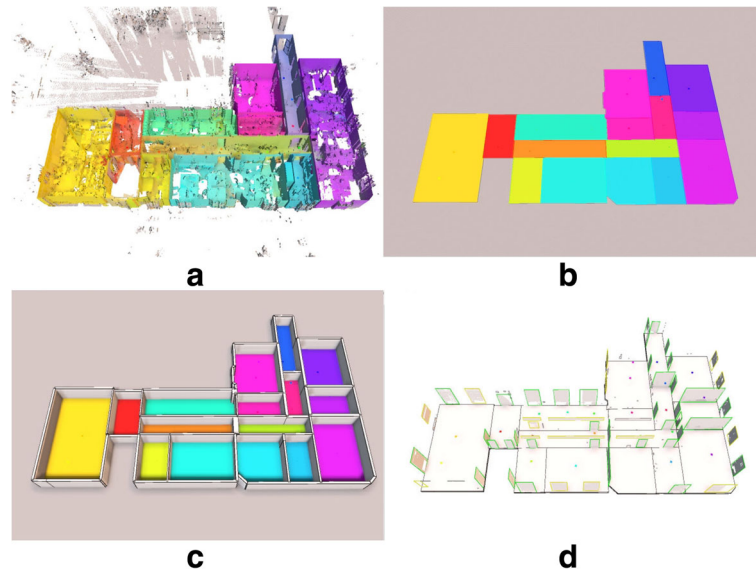
The following sub-chapters describe each part of the pipeline in greater detail.

**Data acquisition**

The input for the pipeline consists of range scan information (point clouds) in E57 format (Huber 2011), accompanied with spherical panoramas, i.e. equirectangular projections of the sphere around the scanner. Bigger datasets, e.g. a whole building floor require several scans and registration. The E57 file stores metadata for each scan; additionally, one panorama per scan has to be supplied. Our test datasets were acquired using a FARO Focus 3D scanner. In some cases the scanner-mounted camera yields too low resolutions for the details that should be detected in photographs, and some lighting conditions were challenging. In these cases, a high-resolution panoramic image was used that was taken with a Canon 500D DSLR and a Nodal Ninja 3 panoramic head. These panoramas were acquired at the scanning position, immediately after the scan.

**Preprocessing**

If multiple scans have been acquired, they need to be registered into one coordinate system before further processing. Registration is usually done using the scanner software and is outside the scope of this paper. After registration, all point clouds and meta data (e.g. position or orientation of each scan) are stored in E57 format. We utilize the approach presented by Ochmann et al. to obtain the room layout and a coarse geometric representation (see Fig. 6): “In contrast to pure surface reconstructions, their approach allows more comprehensive use: first, it enables efficient high-level editing operations in terms of e.g. wall removal or room reshaping which always result in a topologically consistent representation. Second, it enables easy taking of measurements like e.g. determining wall thickness or room areas. These properties render the reconstruction method especially beneficial to architects or engineers for planning renovation or retrofitting. The reconstruction task is cast as a labelling problem which is solved by an energy minimization. This global optimization approach allows for the reconstruction of



**Fig. 6** Geometric Preprocessing. The geometric preprocessing of range scans (a) was carried out using the method of Ochmann et al. (Ochmann et al. 2016). The method estimates a room layout (b), and the configuration of walls (c) and openings, e.g. windows and doors (d)

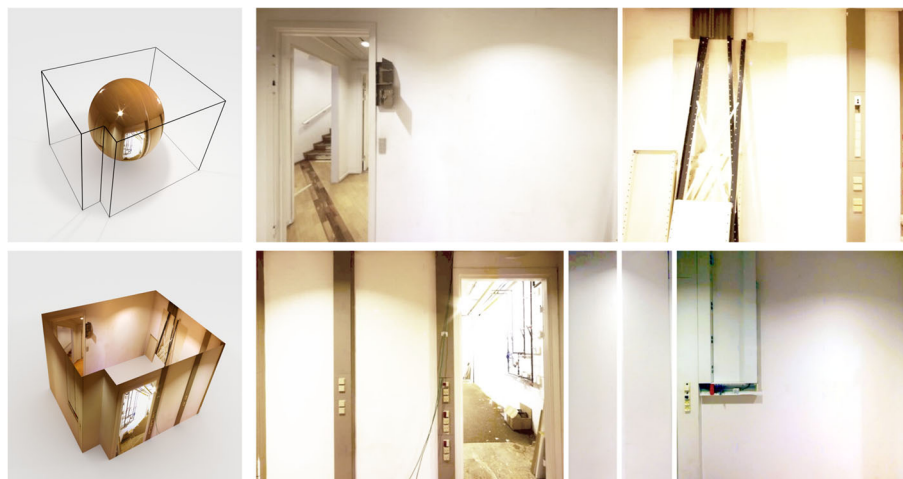
wall elements shared between rooms while simultaneously maintaining plausible connectivity between all wall elements. The reconstructed model is further enriched by detected doors and windows” (Ochmann et al. 2016).

The resulting data is stored in an intermediate JSON format that encodes the position of each wall, the adjacencies of walls, and the position of windows and doors inside walls. After the extraction of rooms and walls, a projection of the spherical panorama onto each wall is generated to obtain orthophotos for each wall using our *orthogen* command line application which is described in more detail in (Krispel et al. 2015). Figure 7 shows an example on a simple room. If several scans are available,

each pixel of an orthophoto is generated using the nearest scan, i.e. the scan with shortest distance from scanning position to the pixel.

**Endpoint detection**

The detection of visible endpoints, e.g. sockets or switches in walls, were carried out using a machine learning approach: the system was trained using a database of images of fixed size, in our case 130 x 130 pixel; the scale of detection was fixed at 1mm per pixel. The learning is carried out using a reduced representation of a training image, the so-called descriptor. For our use case, the descriptor is a combination of a slightly modified



**Fig. 7** Orthophoto generation. From the panoramic views and coarse wall geometry (left), orthophotos for each wall are generated (right)

Histogram of oriented gradients (Dalal and Triggs 2005) and Haar-like features (Viola and Jones 2001). All these descriptors are combined by concatenating their entries to one feature vector. According to its values, the pre-trained random forest classifier (Breiman 2001) retrieves the probability of each class. In order to train the classifier, we created a large training set of labelled image patches representing three object classes: power socket, light switch and background. After classification, a subsequently applied non-maxima suppression on the class probability maps yield the final detection results (Krispel et al. 2015). The results of the detection are bounding boxes around each detected endpoint, associated to a specific orthophoto, or wall, as shown in Fig. 8.

### Electrical line hypothesis

The hypothesis is carried out by two main steps: the installation zone generation and the final hypothesis extraction.

Installation zone generating is carried out by an approach that is inspired by formal grammars (Hopcroft et al. 2006): The generating ruleset encodes domain knowledge from technical standards. Different technical standards are represented by different ruleset files. In the course of the DURAARK project several standards were implemented. All examples in this article use the implementation of the DIN 18015. The information obtained from prior pipeline steps are encoded as starting symbols, e.g. walls or openings, and attributes, e.g. the location of a symbol inside a wall. The evaluation of the generating ruleset will yield horizontal and vertical installation zones per wall. An illustrative simple example is shown in Fig. 9.

After grammar evaluation, a graph is built that encodes all possible routings. A graph consists of a set of elements (vertices or nodes) together with a binary relation that is defined on the set; a relation between two vertices

is called edge (Wilson 1996). Crossing installation zones yield vertices, and detections, as shown by the blue and green rectangles in Fig. 9, are either placed as vertex on an existing edge, or connected by the shortest orthogonal path to nearby edge. Using the adjacency information of the floorplan, adjacent zones of adjacent walls are identified and connected. Furthermore, a power root node, see the orange rectangle in Fig. 9, has to be supplied; in many cases the location of the main fuse box of a floor is known and can be added to the data description, if no root is known one of the endpoints is chosen as root.

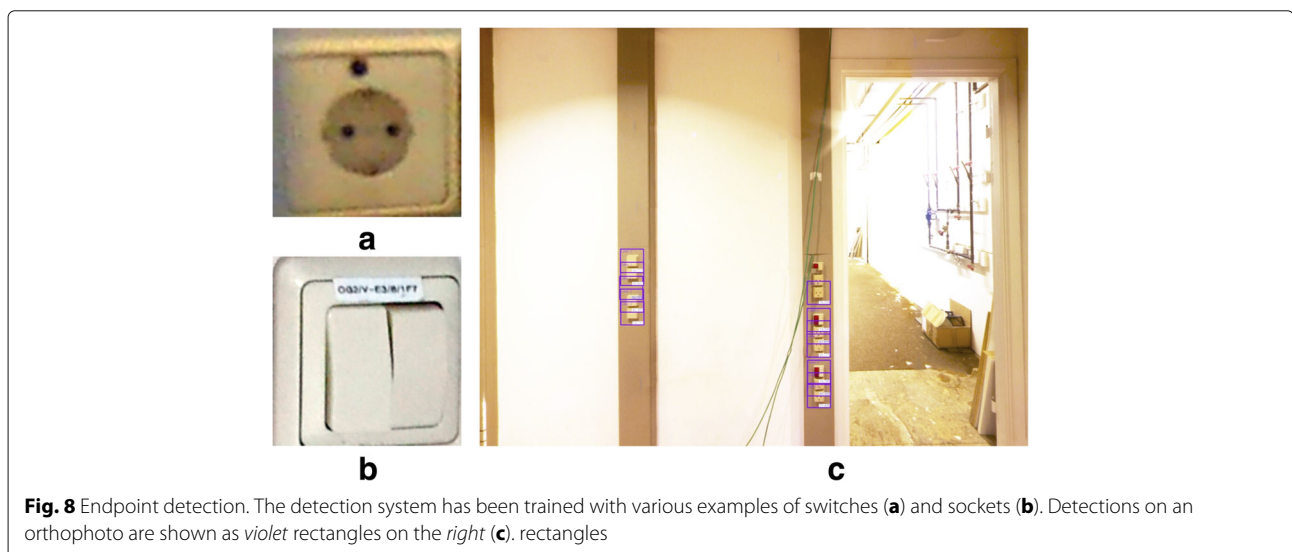
Finally, a subgraph (specifically a forest) is extracted from this graph using a discrete optimization approach. This subgraph connects all endpoints (detections) to a power root, under the assumption that the overall length of power lines, which corresponds to material costs, is minimized. Note that this is not a minimum spanning tree problem (Graham and Hell 1985), as not all vertices of the main graph might be included in the optimal solution.

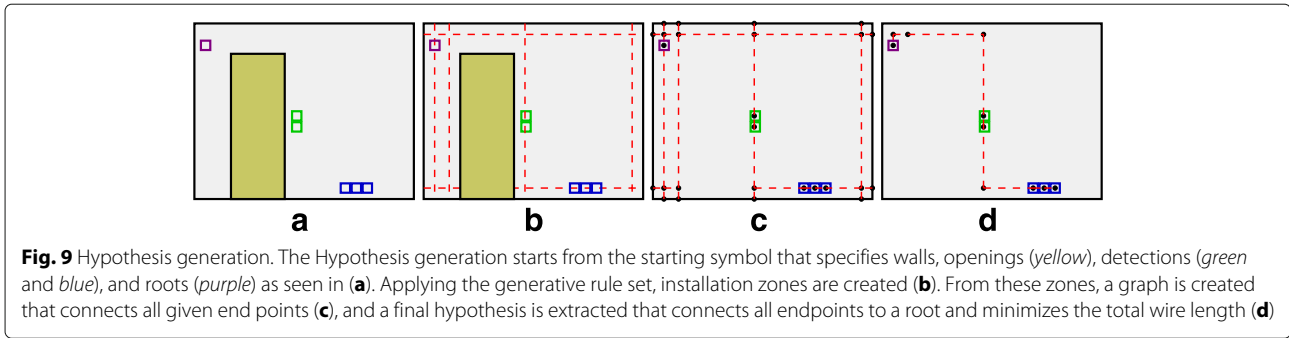
### Installation zone grammar

Installation zones are placed with respect to wall boundaries or “forbidden” zones, e.g. openings such as doors or windows, these objects are utilized as start symbols in the grammar.

We define the installation zone grammar as an attribute grammar  $G_{IZ} = (N, T, P, S)$  that consists of the set of non-terminal symbols  $N$ , the set of terminal symbols  $T$ , the set of production rules  $P$  and the start symbol  $S$ . Nonterminal symbols  $NT \in N$  are written in uppercase letters, e.g. WINDOW and terminals  $t \in T$  are written in lowercase, e.g. hzone. A production rule is written

$$NT \{constraints\} \rightarrow \langle t | NT \rangle \{attribute\ definitions\} \quad (1)$$





The left side of a production rule contains exactly one nonterminal, and an optional constraint that evaluates to a boolean value, the rule will be matched only if the constraint evaluates to true. The right side can consist of any number of nonterminal and terminal symbols, together with attribute definitions in curly brackets. An attribute corresponds to a key-value pair. When a rule is evaluated, all symbols on the right hand side inherit all attributes from the left hand side, attribute definitions on the right hand side are carried out afterwards.

The starting production rule, which is generated from the scanned data, produces nonterminal symbols according to the elements that influence installation zone placement (walls, doors, windows, detections). All positional information is stored in attributes, with respect to a wall coordinate system (see also Fig. 10). In our implementation the wall origin is defined on top left when looking at the wall from inside a room, which corresponds to image coordinates. The terminal symbols of this grammar correspond to horizontal and vertical installation zones, as well as forbidden zones, e.g. doors or windows.

Given such a starting rule, the production rules are evaluated until only terminal symbols are left, and all non-terminal symbols are consumed. The terminals with a special meaning for the hypothesis extraction are *hzone*,

*vzone* and *root*, other terminals are treated as an endpoint if it contains the attribute *endpoint* that is set to *true*. Using this mechanism, users can manually add points that should be contained in the routing hypothesis, e.g. if there are known positions of wirings that could not be detected by the automatic endpoint detection.

This grammar is context free, i.e. there is only exactly one nonterminal on the left hand side of a rule, but it turned out that this approach was not sufficient in cases where there is strong evidence for an installation area that is not given solely by the rules of the technical standard, but when the detected endpoints contain large horizontal or vertical structures, e.g. an array of sockets. Therefore, the grammar description has been extended to also contain rules of the form

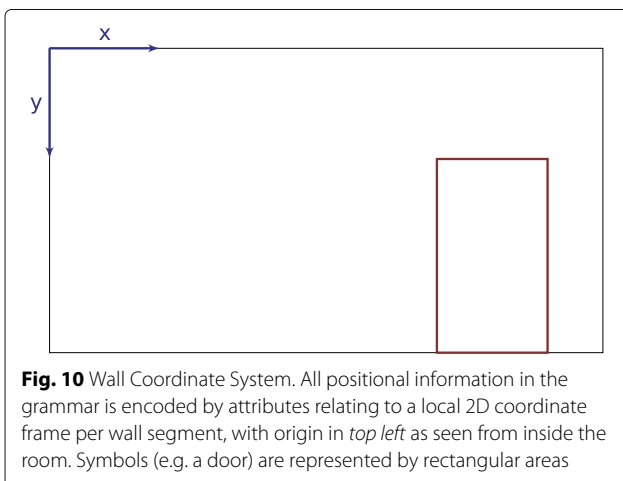
$$NT1 : NT2 \{constraints\} \rightarrow \langle \tau | NT \rangle \{attribute\ definitions\} \tag{2}$$

where the left hand side consists of pairs of nonterminals, written “NT1 : NT2”, with an optional constraint statement. Upon evaluation, the production evaluation system creates for these pair rules all possible pairings of concrete instances of nonterminals. These rules match if their condition statement evaluates to *true*. Pair rules are always prioritized until no pair of nonterminals matches, afterwards context free rules are processed. This allows us to formulate rules that group horizontal and vertical arrangements of endpoints together; groups that exceed a specific size will generate an additional installation zone.

**Wire routing hypothesis**

A wire routing in the proposed method corresponds to a set of connected straight wire segments beneath a wall surface, and is represented by a graph  $G_W = (V, E)$ , that contains vertices  $v \in V$  and edges  $e \in E$ . A vertex is associated to a position in wall coordinates, and an edge connects two vertices  $v_0$  and  $v_1$ . Vertices can also be associated to an endpoint, such as sockets and switches, or a power root, expressed by attributes.

The graph of all possible wire routings is created from the list of terminal symbols after grammar evaluation. A





line arrangement that is formed by the horizontal and vertical installation zones is built. Any edges that intersect forbidden zones, e.g. openings, are removed. Terminals that are marked as endpoints are connected to the graph, distinguished by three different cases: terminals near vertices are directly associated to this vertex, terminals near edges will split the edge and introduce an additional endpoint vertex on the edge, remaining terminals will create a new vertex and connected with either a horizontal or vertical edge to the nearest edge or vertex of the graph. Finally, graphs that correspond to adjacent walls are connected at neighbouring nodes using zero length edges.

The final hypothesis is extracted from this graph under the assumption that the wiring was carried out minimizing material costs. The problem of finding a hypothesis subgraph  $G_H \subseteq G_W$ , where  $G_H = (V_H, E_H)$  is a forest subject to

$$\forall v \in V : v \in V_H \text{ if } v \text{ is endpoint} \quad (3)$$

with the cost function

$$\sum_{v_0, v_1 \text{ of } e \in E_H} d(v_0, v_1) \quad (4)$$

with the penalty  $d$  being the euclidean distance.

This problem is also called the minimum Steiner tree problem in graphs, which is of practical importance in several areas, e.g. chip design or shortest-length connection of households to a power grid. This problem is also known to be NP-complete, as shown by Hwang and Richards (Hwang and Richards 1992), even within an approximation factor of  $\approx 1.129$  which was shown by Kaklamanis et al. (Kaklamanis et al. 2008). For small graphs, e.g. single or few-room datasets, this can be evaluated exhaustively to find the global optimum, for larger datasets a faster, approximative algorithm may be needed. Our current solution implements a local search algorithm that is similar to ‘‘A fast algorithm for steiner trees’’ from Kou et al. (Kou et al. 1981); this algorithm subsequently grows the final graph by sorting endpoints by 3D euclidean distance and connecting each endpoint to the graph.

## Results and discussion

The approach was evaluated on several datasets. Figure 11 shows an example of a single room dataset. The evaluation was carried out on a Core i7-4930K with 3.40 GHz and 6 cores. The orthophoto generation took 2.4s, the endpoint detection 575.2s and the wire hypothesis (inclusive grammar evaluation) 0.44s. The comparatively long runtime of the endpoint detection stems from the exhaustive sliding

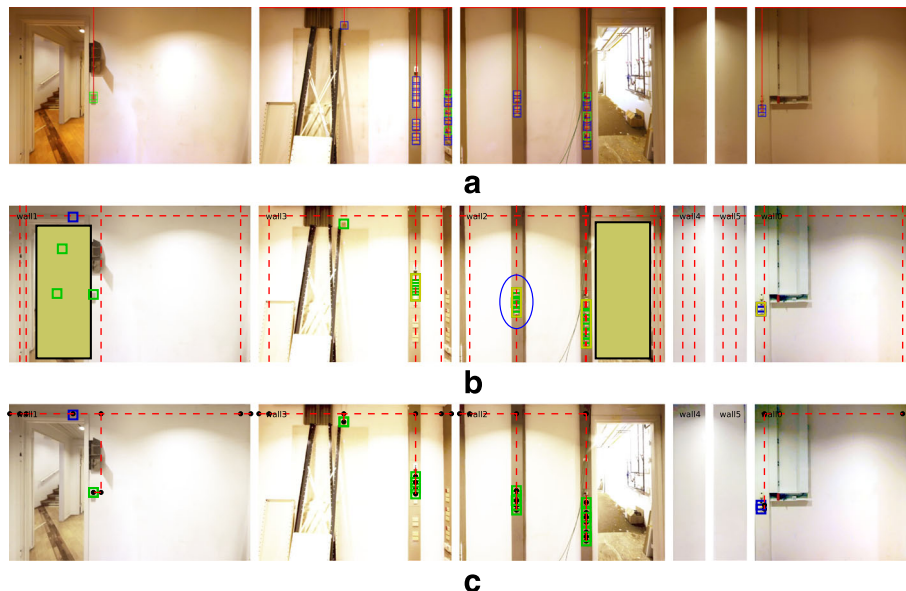
window evaluation of each wall. This could be further improved by using an interest point detector that reduces the search space of endpoint candidates. The detection results are further improved by the pipeline in a step that removes false positives, i.e. an endpoint is wrongly detected, inside forbidden regions, compare also Fig. 11b and Fig. 11c. The grammar that was used places vertical installation zones near doors and windows, and creates zones from vertical or horizontal groupings, as seen in Fig. 11b.

This work presents an approach that synthesizes a hypothesis of electrical lines from range scans and photographs. It uses a new method that utilizes a formal grammar to represent the guidelines that govern the placement of electrical lines.

The main contribution of our new approach is the formalization of norms, standards and prior knowledge in a machine-readable manner. The presented application in the context of building information management shows the potential of the generative paradigm. Unfortunately, the machine-readable representation in form of a grammar system differs significantly from human-readable, descriptive norms. Future research will focus on a representation that closes the gap between machine-readable and human-readable knowledge representations.

Besides scientific contributions the new approach has several benefits: it enables the adaption of rules to different norms, up to specifics that concern only one building. The labeling of detected sockets with confidence values from the detection stage proved to be a viable approach in this project. Concerning future developments practitioners which evaluated our solution remarked that it would be necessary to show users that the generated data is an estimate and not the sole truth, to counter the blind believe, which stakeholders often demonstrate in the face of external data. The representation of an installation zone probability as a heat map would be a way to present this information. Practitioners also stated that this feature would be helpful in the design of new buildings. In later BIM based planning stages, the number and locations of sockets and switches per room are known, but no automated means to estimate the cable length and the attached costs are available.

For a productive usage the following issues have to be addressed. Just like any acquisition and reconstruction method, our approach also relies on the input data quality and completeness (a wall which has not been photographed cannot be analysed); the preprocessing step assumes a single floor with straight walls only; the endpoint detection can only detect what has been learned (in our examples we did not train high voltage sockets); the hypothesis cannot reproduce cases where the actual wiring does not follow the rules – a scenario often found in old buildings. Nevertheless, this



**Fig. 11** Single room dataset evaluation. The first row shows the manually acquired ground truth labeling of sockets and switches (a), as well as electrical lines, according to a domain expert. The second row (b) shows the detected endpoints and the resulting installation zones after grammar evaluation, their center lines are depicted by *dashed red lines*. Note that zones are also generated from context sensitive information from vertically grouped endpoints, e.g. as shown by the ellipse in the middle of wall2 in (b). The final hypothesis is shown in (c)

proof of concept already shows the applicability of our approach.

Source code for the proposed pipeline for the modules orthophoto generation, electric endpoint detection and installation zone generation as well as wiring hypothesis is freely available at the GitHub repository<sup>4</sup> in Modules orthogen, elecddetect and wiregen.

## Endnotes

<sup>1</sup> <http://www.cloudcompare.org/>

<sup>2</sup> <http://x3dgraphics.com/>

<sup>3</sup> <http://www.blender.org/>

<sup>4</sup> <https://github.com/DURAARK/>

## Acknowledgements

Per-Kristian Hanson and Marcus August Frølich Innvær from CITA have been instrumental in the creation of the 3D scan dataset, panoramic images and ground truth for this project. Furthermore, the authors would like to thank Robert Viehauser for the implementation of the computer vision pipeline and Sebastian Ochmann for kindly providing Fig. 6.

## Authors' Contributions

UK is the main author of this manuscript. He is a PhD student at Graz University of Technology and Fraunhofer Austria Research GmbH. He developed the electrical line hypothesis method described in this document. HLE is an Architectural Design Researcher at The Royal Danish Academy of Fine Arts, Schools of Architecture, Design and Conservation CITA, and made key contributions to the data acquisition part for the evaluation of the pipeline. MT is an Associate Professor at CITA and took part in the pipeline design and the Introduction section of this manuscript. TU is Deputy Head of the business area Visual Computing of Fraunhofer Austria Research GmbH and made key contributions to the Knowledge Management section. All authors read and approved the final manuscript.

## Competing interests

The authors declare that they have no competing interests.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Author details

<sup>1</sup>Graz University of Technology & Fraunhofer Austria Research GmbH, Graz, Austria. <sup>2</sup>Centre for Information Technology and Architecture, Copenhagen, Denmark.

Received: 1 November 2016 Accepted: 6 March 2017

Published online: 21 March 2017

## References

- Atanasu, B, Despret, C, Economidou, M, Maio, J, Nolte, I, Rapf, O (2011). *Europe's buildings under the microscope - a Country-by-country review of the energy performance of buildings*. Brussels: Buildings Performance Institute Europe (BPIE). ISBN: 9789491143014.
- Azhar, S (2009). BIM for electrical construction: benefits and current trends. *Journal of Building Information Modeling (JBIM)*, 4, 28–29. The National Institute of Building Sciences buildingSMART alliance. [https://www.wbdg.org/pdfs/jbim\\_fall09.pdf](https://www.wbdg.org/pdfs/jbim_fall09.pdf).
- Bernstein, HM, & Jones, SA (2012). *The business value of BIM in North America*. Bedford: McGraw-Hill Construction.
- Bishop, CM (2007). *Pattern Recognition and Machine Learning. Information Science and Statistics*. New York: Springer.
- Braun, S, Rieck, A, Köhler-Hammer, C (2015). *Digitales Bauen noch Zukunftsmusik*. Stuttgart: Fraunhofer-Institut für Arbeitswirtschaft und Organisation (IAO).
- Breiman, L (2001). Random forests. *Mach Learn*, 45(1), 5–32.
- Dalal, N, & Triggs, B (2005). Histograms of oriented gradients for human detection. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, vol. 1 (pp. 886–893). New York: IEEE.
- European Parliament (2002). Directive 2002/91/EC of the European Parliament and of the Council of 16 December 2002 on the energy performance of buildings. <http://eur-lex.europa.eu/eli/dir/2002/91/oj>. Accessed 11 Feb 2017.

- Graham, RL, & Hell, P (1985). On the history of the minimum spanning tree problem. *Ann Hist Comput*, 7(1), 43–57. doi:10.1109/MAHC.1985.10011.
- Hanna, AS, Yeutter, M, Aoun, DG (2014). State of practice of building information modeling in the electrical construction industry. *J Constr Eng Manag*, 140(12), 05014011. doi:10.1061/(ASCE)CO.1943-7862.0000893. [http://dx.doi.org/10.1061/\(ASCE\)CO.1943-7862.0000893](http://dx.doi.org/10.1061/(ASCE)CO.1943-7862.0000893).
- Havemann, S, Ullrich, T, Fellner, DW (2012). The Meaning of Shape and some Techniques to Extract It. *Multimed Inf Extraction*, 1, 81–98.
- Hopcroft, JE, Motwani, R, Ullman, JD (2006). *Introduction to Automata Theory, Languages, and Computation*, 3rd edn. Harlow: Pearson Education Limited.
- Huber, D (2011). The ASTM E57 File Format for 3D Imaging Data Exchange, In *Proceedings of the SPIE Vol. 7864A, Electronics Imaging Science and Technology Conference (IS&T), 3D Imaging Metrology*. Copyright 2011 SPIE and IS&T. One print or electronic copy may be made for personal use only. Systematic reproduction and distribution, duplication of any material in this paper for a fee or for commercial purposes, or modification of the content of the paper are prohibited, vol. 7864A. Bellingham: Society of Photographic Instrumentation Engineers (SPIE).
- Hwang, FK, & Richards, DS (1992). Steiner tree problems. *Networks*, 22(1), 55–89. doi:10.1002/net.3230220105.
- ISO (2013). Industry Foundation Classes (IFC) for data sharing in the construction and facility management industries. Technical Report 16739:2013, International Organization for Standardization.
- Kaklamanis, C, Chlebič, M, Chlebičová, J (2008). Algorithmic aspects of global computing the steiner tree problem on graphs: Inapproximability results. *Theor Comput Sci*, 406(3), 207–214. doi:10.1016/j.tcs.2008.06.046.
- Krispel, U, Evers, HL, Tamke, M, Viehauser, R, Fellner, DW (2015). Automatic texture and orthophoto generation from registered panoramic views. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XL-5/W4*, 131–137. doi:10.5194/isprsarchives-XL-5-W4-131-2015.
- Krispel, U, Schinko, C, Ullrich, T (2015). A survey of algorithmic shapes. *Remote Sens*, 7(10), 12763. doi:10.3390/rs71012763.
- Kou, L, Markowsky, G, Berman, L (1981). A fast algorithm for steiner trees. *Acta Informatica*, 15(2), 141–145. doi:10.1007/BF00288961.
- Malleson, A (2015). NBS National BIM Report 2015, In *NBS*. <https://www.thenbs.com/knowledge/nbs-national-bim-report-2015>. Accessed 11 Feb 2017. [www.theNBS.com](http://www.theNBS.com).
- OAS/USDE. Drawings to Accompany the Building Guidelines, Section G: Electrical Guidelines. <http://www.oas.org/pgdm/document/codedraw/sectiong.htm>. Accessed 11 Feb 2015.
- Øbro, M (2015). *BIM Survey 2014*. Herlev: Byggeri Informationsteknologi Produktivitet Samarbejde (BIPS). [http://bips.dk/files/news\\_files/bim\\_survey\\_2014\\_r0.pdf](http://bips.dk/files/news_files/bim_survey_2014_r0.pdf).
- Ochmann, S, Vock, R, Wessel, R, Klein, R (2016). Automatic reconstruction of parametric building models from indoor point clouds. *Comput & Graph*, 54, 94–103. doi:10.1016/j.cag.2015.07.008. Special Issue on CAD/Graphics 2015.
- Petkova, PJ, & Rüppel, U (2014). A graph-based prediction method for electrical wiring in old residential buildings as a part of bim for urban mining purposes. *eWork and eBusiness in Architecture, Engineering and Construction*, 109–113. doi:10.1201/b17396-21.
- Riemenschneider, H, Krispel, U, Thaller, W, Donoser, M, Havemann, S, Fellner, DW, Bischof, H (2012). Irregular lattices for complex shape grammar facade parsing. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 25, 1640–1647.
- Rapf, O, Faber, M, Marian, C, Fata, FA (2015). *Renovating Germany's Building Stock*. Brussels: Buildings Performance Institute Europe (BPIE). ISBN: 9789491143137.
- Schinko, C, Strobl, M, Ullrich, T, Fellner, DW (2010). Modeling Procedural Knowledge – a generative modeler for cultural heritage. *Proc EUROMED 2010 - Lect Notes on Comput Sci*, 6436, 153–165.
- Tamke, M, Evers, HL, Zwierzycki, M, Ochmann, S, Vock, R, Klein, R (2016). An automated approach to the generation of structured building information models from unstructured point cloud scans. In K Kawaguchi, M Ohsaki, T Takeuchi (Eds.), *Proceedings of the IASS Annual Symposium 2016 "Spatial Structures in the 21<sup>st</sup> Century"*. Madrid: The International Association for Shell and Spatial Structures.
- Thaller, W, Krispel, U, Zmugg, R, Havemann, S, Fellner, DW (2013). Shape Grammars on Convex Polyhedra. *Comput & Graph*, 37, 707–717.
- Thaller, W, Zmugg, R, Krispel, U, Posch, M, Havemann, S, Fellner, DW (2013). Creating procedural window building blocks using the generative fact labeling method. *ISPRS - Int Arch Photogramm Remote Sens Spat Inf Sci, XL-5/W1*, 235–242. doi:10.5194/isprsarchives-XL-5-W1-235-2013.
- Vækstministeriet, E (2006). Stærkstrømsbekendtgørelsen, EK nr 9146 af 22/03/2006 Gældende (SB - Afsnit 6A) Offentliggørelsesdato. <https://www.retsinformation.dk/forms/R0710.aspx?id=27280>. Accessed 11 Feb 2017.
- Viola, P, & Jones, M (2001). Rapid object detection using a boosted cascade of simple features, In *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference On*, vol. 1 (pp. 1–511): IEEE.
- Wikimedia Commons (2014). Verlegezonen Wohnräume. [https://commons.wikimedia.org/wiki/File:Verlegezonen\\_Wohnr%C3%A4ume.jpg](https://commons.wikimedia.org/wiki/File:Verlegezonen_Wohnr%C3%A4ume.jpg). Accessed 11 Feb 2017.
- Wilson, RJ (1996). *Introduction to Graph Theory*, 4th edn. Harlow: Pearson Education Limited.

Submit your manuscript to a SpringerOpen® journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](http://springeropen.com)