ROOM SHAPES AND FUNCTIONAL USES PREDICTED FROM SPARSE DATA

Y. Dehbi[a]* , N. Gojayeva[a], A. Pickert[a], J-H Haunert[a] and L. Plümer[a,b]

[a] Institute of Geodesy and Geoinformation, University of Bonn, Meckenheimer Allee 172, Bonn, Germany -
(dehbi, pluemer, haunert)@igg.uni-bonn.de
(s7nagoja, s6anpick)@uni-bonn.de

[b] Southwest Jiaotong University, Chengdu, China

Commission IV, WG IV/5

KEY WORDS: Indoor Model, CityGML, BIM, Bayesian Classification, Stochastic Reasoning

ABSTRACT:

Indoor models are highly relevant for indoor navigation. However, they are hard to achieve if high-resolution data is not available. Many researchers used expensive 3D laser scanning techniques to derive indoor models. Few papers describe the derivation of indoor models based on sparse data such as footprints. They assume that floorplans and rooms are rather rectangular and that information on functional use is given. This paper addresses the automatic learning of a classifier which predicts the functional use of housing rooms. The classification is based on features which are widely available such as room areas and orientation. These features are extracted from an extensive database of annotated rooms. A Bayesian classifier is applied which delivers probabilities of competing class hypotheses. In a second step, functional uses are used to predict the shape of the rooms in a further classification.

1 MOTIVATION AND CONTEXT

Beyond the visualization tasks, 3D city models are nowadays a prerequisite for a wide range of other applications. Building models, in particular, are employed for the automatic estimation of energy demand, solar irradiation or the improvement of positioning in urban environments. This significant progress is attributed to the technological advances in wide-range mapping, for instance using mobile platforms, and the automatic reconstruction and interpretation of the acquired data. Biljecki et al. (2015) give an overview on the applications of 3D city models.

Likewise, the last decade is characterised by an increasing demand of 3D indoor models. The latter are mainly used for indoor navigation, facility management and monitoring of indoor environments (Zlatanova and Isikdag, 2015). The automatic derivation of indoor models is, however, still a challenging task. In contrast to the exterior of buildings, researchers are facing many difficulties, e.g. weak GPS signal and occlusions caused by furnitures, in order to enable an automatic acquisition of useful indoor data. For this reason, few approaches tried to generate indoor models with as few observations as possible in order to overcome the mentioned data capturing problems (Loch-Dehbi et al., 2017; Rosser et al., 2017; Dehbi et al., 2017).

The methods demonstrated by Dehbi et al. (2017) and Loch-Dehbi et al. (2017) deliver statistically qualified floorplan hypotheses based on few observations. These observations consist of information such as room areas, functional use, window locations and, possibly, room numbers for office rooms. The room numbers are exploited for deriving knowledge about room adjacencies of office rooms. These observations together with background knowledge consisting of probability density functions of shape and location parameters of rooms are mostly sufficient for the prediction and derivation of indoor footprint layouts.

As yet, the layout predictions assumed rectangular rooms and relied on numbered rooms. Numbering, however, is missing in
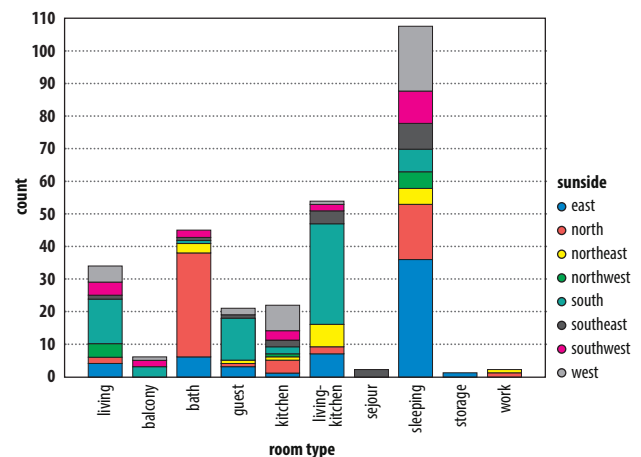

Figure 1. Functional use and orientation of housing rooms in Germany – People prefer to spend dinner in a sunny living room.

housing rooms which increases the degree of complexity of the performed stochastic and geometric reasoning. Furthermore, housing rooms are characterised by a high shape variety. In this paper, we demonstrate that classification methods are appropriate to derive the knowledge about functional use as well as shape categories of rooms. We hypothesize that the results of our classification are highly relevant for the extension of the indoor reasoning approaches based on sparse observations and, hence, improve the prediction quality of the derived housing indoor models. This represents a contribution to establish the missing link between CityGML (Gröger and Plümer, 2012) and BIM models.

In this paper, we demonstrate that we are able to predict the functional use and shape of rooms without the need of additional measurements. The predictions are qualified with probabilities enabling the assessment and comparison of alternative models. The incorporation of statistical knowledge, inferred from learning examples, gives insight into layout preferences characterising

_____
*Corresponding author

the design of building rooms. For instance, the information about sun side turns out to be a good feature for the identification of the functional use of rooms. Figure 1 visualises the room orientation preferences of German people and architects with regard to the functional use of each room. Especially living rooms in Germany are directed in such a way that the sun from the south can be exploited. This aspect, however, depends on the geographic region taken into consideration. In this context, our data analysis reveals other layout preferences in Morocco with regard to the sun side. In this sense, the sun side feature is playing a significant role not only for the functional use identification but also for the positioning of a room within the housing footprint.

The approach of Loch-Dehbi et al. (2017) is able to deal with non-rectangular footprints; it assumes, however, rooms to be rectangular. Since we used a probabilistic classifier, Naive-Bayes, which outputs qualified class predictions, we performed a second classification based on the predicted functional use. This step consists in predicting the shape of rooms in order to extend the layout derivation by dealing with non-rectangular rooms.

All in all, both classification results can be used to support predictions in reasoning approaches based on sparse observations. Following the same paradigm, we designed features based on information which can be easily extracted from facility management service data or observed from the outside of the considered house. The combinatorial part for predicting housing rooms can be augmented by probabilistic background knowledge, in particular if the underlying architectural constraints are weak. This is for example the case for exterior walls with few windows.

The remainder of this paper is structured as follows. The next section discusses related work. Section 3 gives insight into the classification tasks that are the subject of this paper and their results. Especially section 3.1 deals with the prediction of functional use based on weak observations, whereas 3.2 explains how the shape category of rooms can be inferred based on the predicted functional use. Both subsections discuss the experimental results. The paper is summarized and concluded in Section 4.

2 RELATED WORK AND CONTRIBUTION

Indoor models are nowadays gaining more and more attention. They are of high relevance for various applications such as guide for the blind, rescue management and evacuation planning. In the context of BIM, they are used during the planning and construction stage and for facility management later on. Wenming et al. (2018) proposed an approach for the layout design for building interiors based on mixed integer quadratic programming. This approach is however rather suitable for the generation of indoor models for buildings in planning and does not address as-built models. The derivation of as-built indoor models from surveying is however a still challenging task. In this context, Turner and Zakhor (2014) triangulated a 2D sampling of wall positions in order to generate building floorplans from laser range data.

Peter (2017) presented a modelling method of indoor environments for the support of reconstruction tasks using Lindenmayer systems. In the same context, formal grammars have been employed by Becker et al. (2015) for the reconstruction of 3D indoor models from 3D point clouds. In another grammar-based approach, Yue et al. (2012) determined the interior layout of buildings based on a shape grammar which describes different building styles. Following the same spirit, shape grammars have been used by Khoshelham and Díaz-Vilariño (2014) for 3D indoor modelling using an iterative placing, connecting and merging of
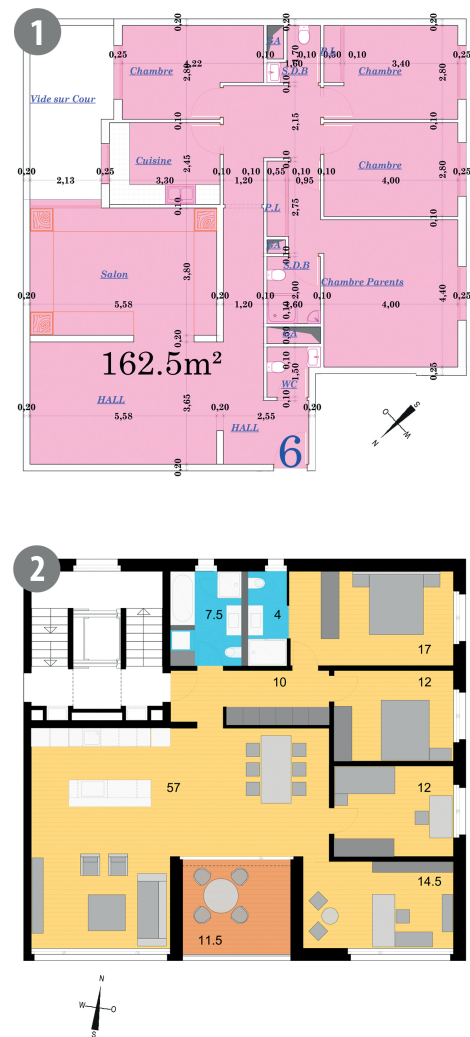

Figure 2. Two examples of openly accessible floorplans used as training data in our experiments. Different geographic regions (1:Morocco, 2: Germany/ Switzerland), different floorplans and room shapes (e.g. L and I) are considered.

cuboid shapes. All these grammar-based approaches require a pre-design of an according set of grammar rules.

For scene reconstruction, Ochmann et al. (2016) segmented a point cloud into rooms and outside area using an energy minimization in a labelling problem. Mura et al. (2014) presented an approach for reconstructing the main architectural structure of complex indoor environments given a set of cluttered 3D input range scans. All mentioned approaches require dense observations such as 3D point clouds from laserscans or range cameras using mobile mapping systems that are both cost and time expensive.

Several approaches tried to avoid the measurement overhead, and hence overcome the costly data capture by using low cost sensors. To this aim, Rosser et al. (2015) constructed as-built plans of residential building interiors. In this constraint-based approach, mobile phone sensor data is used for the prediction of interior models. Mobile devices have been employed by Pintore et al. (2016) to capture images for the generation of 2.5D indoor maps. Diakité and Zlatanova (2016) used the low cost Android tablet from Google's Tango project for the acquisition of indoor building environments.
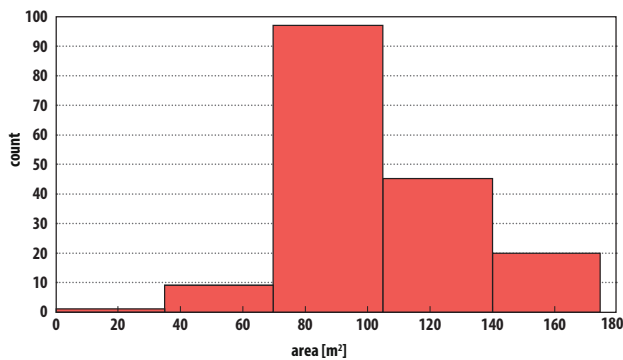
Figure 3. Housing areas which have been considered in our experiments.

While most approaches require measurements of high density such as 3D point clouds or images, Loch-Dehbi et al. (2017) proposed an approach which gets along with few observations in order to predict floorplans of high accuracy. They estimated floorplan parameters and performed a model selection for probabilistically qualified floorplan hypotheses. The sparse observations like the area of rooms and footprints are used together with strong model assumptions supported by a profound background knowledge leading in a two-staged, combinatorial and stochastic, reasoning process. In the same context, based on limited prior knowledge, Rosser et al. (2017) presented a semi automatic data-driven estimation of 2D building interior floorplans. Following the same spirit, we propose an approach which does not rely on dense observations for supporting the prediction of indoor models based on sparse observations. Our aim is to provide key information such as the functional use and the shape of housing rooms without the need of exhaustive measurements. We demonstrate that classification techniques are suitable to acquire such information for supporting, for instance, the combinatorial part in the method of Loch-Dehbi et al. (2017).

3 PREDICTION OF FUNCTIONAL USE AND SHAPE OF ROOMS

The reasoning approach described by Loch-Dehbi et al. (2017) consisted of two components: a combinatorial component and a stochastic component. The approach makes some architectural assumptions and uses thresholds derived from probability density functions (pdfs) of each model parameter. We believe that an extension of the former approach with different room types and housing rooms without room numbers for the explicit addressing of uncertainty may be beneficial. Our goal is to demonstrate that classification methods are suitable to gain knowledge about functional use as well as shape categories of rooms. In this context, the information derived by the classifier is valuable and of great interest.

In this paper, our aim is to learn prediction models for both functional uses and shapes of rooms. The learning process is performed in a supervised way. Thus, we prepared learning examples beforehand. In this context, we used about 1800 labelled rooms annotated within a spatial relational database with several room types. The database is designed in such a way that relevant features are a-priori modelled in order to capture their values stemming from an annotation process. In this context, we collected openly accessible floorplans from different platforms. Figure 2 shows some examples of floorplans which have been used to build labelled training data for our classification. Herewith, we took floorplans from different geographic areas (1: Morocco and

2: Germany/ Switzerland) into consideration. The chosen floorplans are also characterised by different room shapes, e.g. I or L. Likewise, the housing rooms themselves are characterised by a shape and type variety. We selected scaled floorplans in order to perform geometric measurements and extract geometric features accordingly. The directional information with regard to the sun side is mostly available (cf. Figure 2), otherwise we make use of `GoogleEarth` [1] or `OpenStreetMap` (OSM) [2] in order to enrich our rooms with the missing sun side information. Figure 4 illustrates the schema of our spatial relational database. The eight possible cardinal directions are encoded and stored in the table `direction`. This table is in turn referencing the table `window` capturing the directional information of a corresponding room in the table `room`.

We took a particular care to ensure a representative set of learning data. In this context, we annotated floorplans with different range of areas. Figure 3 gives insight into the distribution of the areas of the considered housing floorplans. It can be stated that flats with areas between 70 and 100 m^2 are common. Beside taking different geographic areas (Morocco, Germany/ Switzerland) into account, the buildings and housing rooms in a specific area, especially those from Germany/ Switzerland, stemmed from different construction eras and have different building styles and types.

This section presents two classification approaches in order to predict the functional use of rooms and their shapes based on weak observations. In this context, subsection 3.1 is dealing with the classification of the functional uses, whereas subsection 3.2 gives insight into the classification approach of room shapes based on the a-priori known functional use acquired from the previous classification task.

3.1 Classification of room functional uses

The functional use of rooms in buildings and housing rooms in particular is playing a prominent role for the design and the layout of the according floorplan. Lacking of such information makes it difficult to decide about the arrangement of rooms within a floorplan. This section demonstrates that the prediction of the functional use of rooms within a housing floorplan is possible without the need of exhaustive measurements. A set of relevant features for a particular functional use is designed, analysed and used to train a classifier based on a-priori labelled examples. Figure 5 illustrates the considered functional uses of rooms. Especially, sleeping rooms, living rooms, bath rooms and kitchens are considered. In particular, we are discriminating between ordinary living rooms and those combined with a kitchen. This type will be denoted as `living-kitchen` in this paper. The types of functional use are categorised in the table `functionallUse` according to the database schema (cf. Figure 4).

Man-made objects are for many reasons, such as economical or aesthetic, often characterised by architectural regular patterns. Hence, the identification of such regularities, and modelling them via a sound feature set is a key step towards learning and predicting a given class of interest. For example, the choice of the orientation is a crucial step during the construction planning phase of a building in order to exploit the sun in an optimal way. This economises not only heating and electricity costs but also has impact on the well-being in the housing rooms. A data analysis based on the captured data is confirming these architectural intentions in Germany as already shown in Figure 1. In this context, `sleeping rooms` are mainly exposed to the east in order to enjoy the morning sun. Since `bath rooms` are not relying on

[1] `https://www.google./maps`
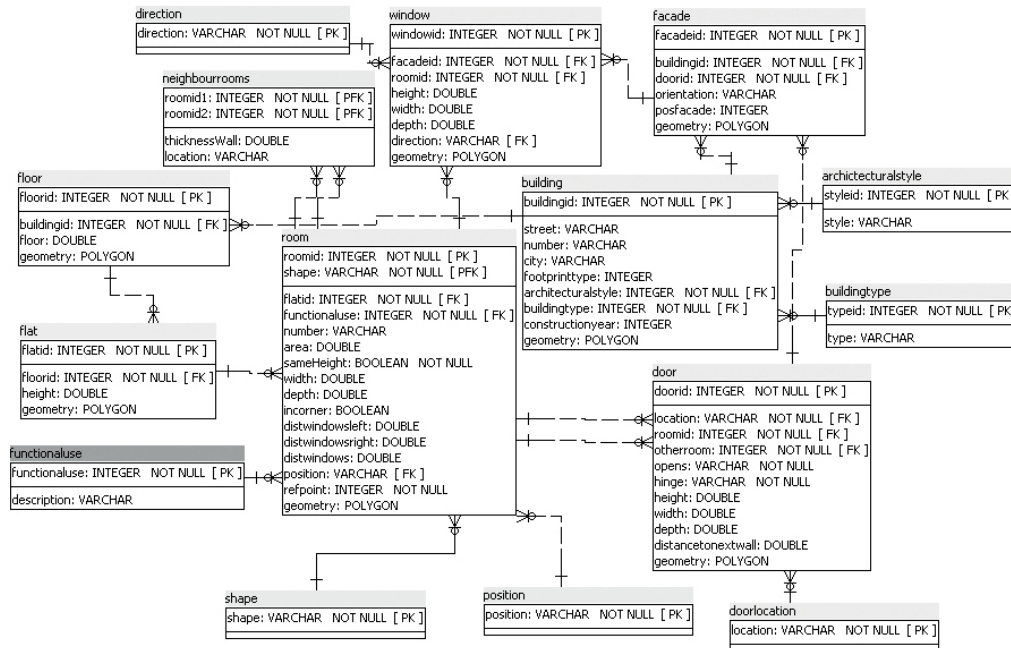[2] `https://www.openstreetmap.org`

Figure 4. Excerpt from the spatial relational database schema. Features are extracted from this data base.

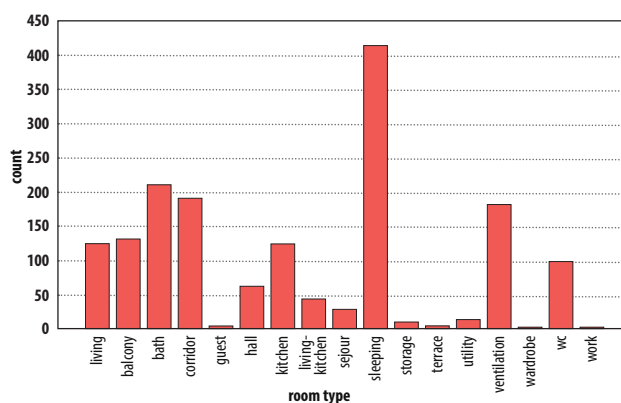Figure 5. A histogram of the different functional uses of rooms taken into consideration in our experiments.

Figure 6. The distribution of functional uses with regard to sun side in Morocco

daylight, they are usually directed along the north. For kitchens, there is no specific pattern to induce. Except for that we are observing a slight dominance of the west direction which is characterised by an evening sun in all seasons. This statement is changing, once the kitchen not only serves as cooking room but also as living room simultaneously which characterises combined `living-kitchen` rooms. In this case, the rooms are mostly directed to the south side. Likewise, balconies and living rooms are also exposed to the south sunlight. Obviously, the sun side factor is depending on the considered geographic region. That is why we analysed annotated data stemming from a different region, namely from Morocco. This analysis revealed other preferences of room placements with regard to the sun side as depicted in Figure 6. For `sleeping` rooms for example, there is a slightly dominance of the south east direction compared to Germany where the east direction is preferred.

The first step in our learning process consists in exploiting the annotated floorplans and rooms. Hence, for each room type a feature set is prepared and extracted from the underlying spatial relational database. Following the spirit of Loch-Dehbi et al. (2017), our aim is to infer and predict the functional use based on
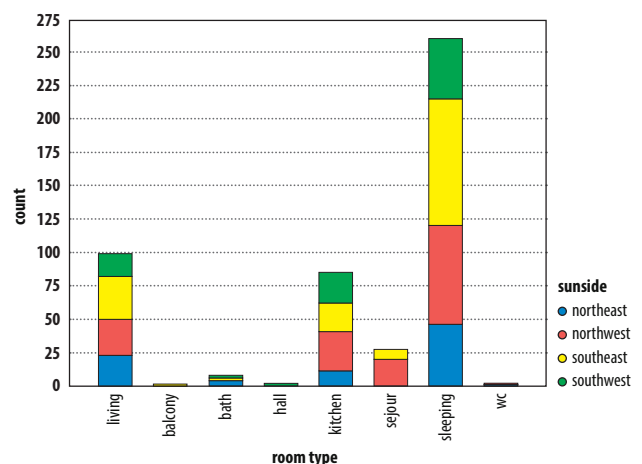
as few and easily accessible sparse observations as possible. In this sense, we are focusing on learning a prediction model based on information which can be easily extracted from facility management service data or observed from the outside of the house of interest. In this context, we designed a feature vector consisting, among others, of information such as whether a room is a corner room or not. As mentioned, the room direction information seems to be a good feature which the functional use depends on. This information is derived for each room having at least one window from the table `window` (cf. Figure 4). Furthermore, we investigated how important the proportional area of a given room is with respect to the whole footprint area for the categorisation of the functional use. For training issue, this geometric feature is derived from the table `room`. For a better discrimination, we discretize areas into ranges using a categorical feature `categorised area` instead of real-valued areas.

In order to analyse the impact of the reduction of the feature vector dimension on the learning results, we performed a feature selection. Table 1 shows an excerpt of the most important used

features for the prediction of the class `living-kitchen` rooms. The corresponding weight of each feature in Table 1 has been automatically derived in the training phase using the *Relieff* algorithm for feature weighting (Robnik-Šikonja and Kononenko, 2003). This weight reflects the importance of each feature with regard to the given class. *Relieff* ranks individual features according to their relevance in the context of others. In this sense, features with low weights are neglected. *Relieff* estimates the quality of neighboured features with regard to their contribution to distinguish between instances of the same and different classes. The feature relevance is measured by sampling instances of the training data and comparing the value of the feature of interest with the according value of the nearest example from the same and a different class. We manually tested several values for the amount of neighbours necessary to calculate the importance for each feature, and report the results for $k_{\mathrm{reliefF}} = 30$. Built upon the derived weights, we ordered and selected the top 90% features with the highest weights. Please note that the weights in Table 1 represent ranked weights and are not probabilities which sum to 1. We normalized numerical features using a z-transformation to zero mean and unit standard deviation to ensure an equal weighting of each single feature.

top features for the class `living-kitchen` room	weight
categorized area	1
proportional area	0.41
direction	0.23
is corner room	0.15
has balcony	0.1
number of windows	0.07
....	...

Table 1. Excerpt of the most important features and their weights for the classification of the functional use `kitchen-living`. The weights are derived using the `reliefF` algorithm.

Generally, our task is to classify a vector of discrete-valued features, $x \in 1 \dots k^D$. Herewith, k denotes the number of values for each feature, and D is the number of features. To this aim, we used a generative approach which specifies the class conditional distribution $p(x|y=c)$.

$$p(x|y=c,\theta) = \prod_{j=1}^{D} p(x_j|y=c,\theta_{jc}), \qquad (1)$$

where y is a label for a given class c.

This approach is assuming that the features are conditionally independent. Although this assumption does usually not hold, the classification results, however, turn out to be good in practice (Bishop, 2007). Depending on the feature property, the class-conditional density can take the following forms:

- real-valued features: Gaussian distribution
  $p(x|y=c,\theta) = \prod_{j=1}^{D} N(x_j|\mu_{jc}, \sigma_{jc}^2)$.

- binary features, $x_j \in \{0,1\}$: Bernoulli distribution
  $p(x|y=c,\theta) = \prod_{j=1}^{D} Ber(x_j|\mu_{jc})$, where μ_{jc} denotes the probability that feature j belongs to the class c.

- categorical features, $x_j \in \{1 \dots k\}$: Multinoulli distribution
  $p(x|y=c,\theta) = \prod_{j=1}^{D} Cat(x_j|\mu_{jc})$, where μ_{jc} denotes a histogram over the k possible values for x_j in the class c.

	true liv-kit	true non liv-kit	precision (%)
pred. liv-kit	48	2	96
pred. non liv-kit	1	47	97.92
recall (%)	97.96	95.92	

Table 2. Classification results of `combined living-kitchen` (liv-kit) rooms; accuracies derived by 10-fold cross validation.

For unknown instances, the goal is to predict the belonging class label y:

$$p(x|y=c,\theta) = \prod_{j=1}^{D} p(x_j|y=c,\theta_{jc}), . \qquad (2)$$

For instance, the functional use types in Figure 4 derived from the table `functionlUse` are used as class labels for the training data in our experiment and have to be inferred for unseen instances later on. For more details, the interested reader is referred to Bishop (2007).

Since we are interested in a stochastic model delivering posterior probabilities of the predicted class, we make use of a generative model consisting in `Naive-Bayes` classifier (Murphy, 2012) as described above. In order to avoid distribution assumptions, especially on real-valued features, we applied a kernel based Bayesian network classification approach (John and Langley, 1995; Prez et al., 2009). In this context, a non-parametric kernel density estimation is used to model the conditional density of continuous variables (Scholkopf and Smola, 2001; Wand and Jones, 1994). Beside omitting parametric assumptions, this has further the advantage that loss of information is prevented after a discretization of continuous features such as the area of rooms.

Table 2 presents the result of the classification with an accuracy of 96.89% and a precision of 97.92% for combined `living-kitchen` rooms derived by a 10-fold cross validation. Further, during the learning of the model the problem of unbalanced class frequencies has been taken into account. In practice, the number of positive examples is usually higher than the number of negative labelled samples. For this reason, we extracted equally sized subsets of positive and negative examples from originally about 512 labelled samples of the category `living-kitchen` rooms. This improved the class precision of this category from about 75% up to 97%. The reported results are achieved using a Bayesian classification with a Gaussian kernel.

Figure 7 confirms the results of the feature selection and illustrates the importance of the feature `proportional area` for the prediction of the class `living-kitchen` room. The room areas are categorised into five categories (A-E) in descending order. Category A characterising rooms with an area from 20 m^2 upwards turns to be a good prior to identify `living-kitchen` rooms. According to our labelled examples, the average area of such rooms is 38 m^2.

Likewise, we designed and implemented a classifier for each room functional type. For example, an accuracy of 88.22% with a precision of 88.43% has been achieved for sleeping rooms. For balconies, the accuracy and the precision amount to 93% and 91.67% respectively. For the implementation of our predictive models and experimental prototyping, we used `RapidMiner`[3] as one of the well-known and widely used machine learning environments.
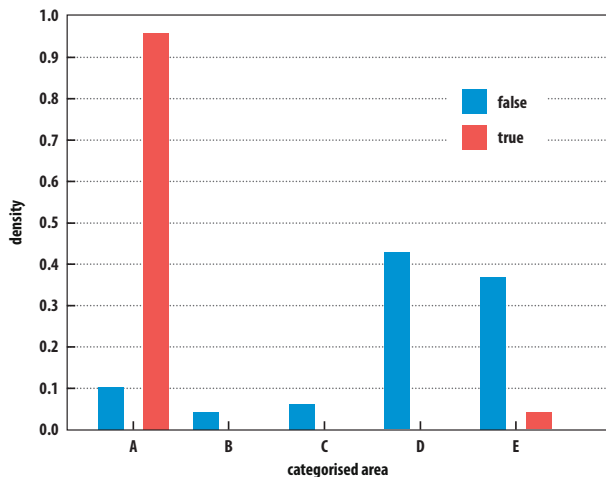
---

[3] https://rapidminer.com/products/studio/

Figure 7. The impact of the feature categorised area on the classification of `living-kitchen` rooms (true). A room in category A is a `living-kitchen` room with a high probability.

3.2 Classification of room shapes

Like functional uses, room shapes in housings have also a massive impact on the layout design of floorplans. Hence, the acquisition of the shape information leads to a significant restriction of the search space of possible room arrangements within an underlying footprint. To this aim, the results of the prediction of functional use achieved in section 3.1 can be exploited as prior knowledge in order to predict room shapes. In this sense, we assumed that the functional use of a given room is known, and the task now is to learn a classification model to predict its shape. In this context, we are discriminating between four types of room shapes as depicted in Figure 8. We can observe that I-shaped rooms (rectangular rooms) are dominating the other shape categories. The second important class is the L-shaped rooms. The two other classes, namely `trapezoidal` and Z are relatively seldom. Since the number of trapezoidal and Z-shaped rooms is not high enough for a good classification, we focused our experiments on the discrimination between I- and L-shaped rooms.

In the same way as described in section 3.1, we applied supervised learning techniques for the prediction of room shapes. To this aim, we make use of the labelled examples from our spatial relational database. In order to avoid unbalanced class frequencies, we prepared 600 learning examples (300 per class). Hence, our task is to train a classifier which discriminates between `L-shaped` and `I-shaped` rooms. The label of each room is extracted from the Table `shape` (cf. Figure 4). Figure 9 shows the distribution of functional use with regard to room shapes. This enables us to get an impression of the influence of the functional use of a room on its shape. It can be stated that the class `living-kitchen` rooms are mainly L-shaped. The I-shape is dominating the remaining shapes, however, the proportion of L-shaped `bath` rooms and `corridors` is non-negligible.

In the same manner as in section 3.1, we performed a feature selection based on `ReliefF` algorithm. This step reveals that the information about the functional use acquired from the previous classification is an important feature for the classification of room shapes. Table 3 shows the most relevant features for this classification. As before, area specific features (categorised and proportional area) are also meaningful for the shape classification. Further, the number of rooms as well as the direction are of great relevance. Both features `numberOfRooms` and `hasBalcony` are influencing the room shapes. The latter describes whether a room
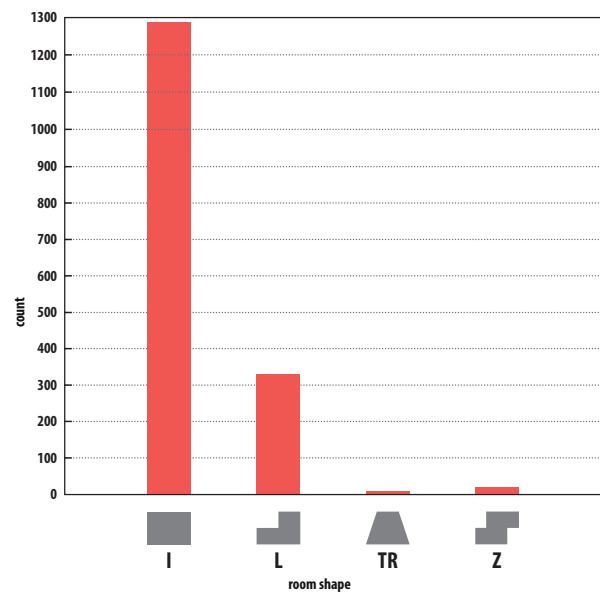

Figure 8. Underlying room shapes in our spatial database. I-, L-, TR- (Trapezoidal) and Z-shaped rooms have been considered.

is connected to a balcony or not. This feature is extracted using the tables `neighbourrooms`, `room` and `fuctionalUse` (cf. Figure 4). One further considered feature encodes the information whether a room is a corner room or not. This turns to be a good feature for the classification of the shape as well as previously for the functional use of a room. The shape of the footprint is also influencing the room shape accordingly. This feature is appearing in the top 10 features with a weight of 0.15.

Beside the mentioned features, the feature vector is augmented by an information about the rough position of an underlying room. This is performed based on discretization of the given footprint. A room position is then encoded depending on common sides shared with the exterior walls. The position is varying depending on the shape of the belonging floorplan. Figure **??** shows the possible positions for two different room shapes. The floorplan (A) is I-shaped, and it is characterised by 9 possible room positions. Exterior walls are highlighted in grey coloured hose. The room with the position 1 has two exterior walls, namely wall 1 and wall 2. The floorplan (B) has an L-shaped boundary. The possible position encodings for such a layout are listed.

Table 4 depicts the classification results of the room shapes. An accuracy of 80.17% has been achieved with a precision of 81.31% for the I-shaped rooms. In comparison to this class, we achieved a lower precision for L-shaped rooms. The classification has been performed using a 10-fold cross validation based on a kernelized Bayesian classifier. It should be noted that the floorplans have been solely used to extract features to build our training examples during the learning phase. For the prediction of the class label of new instances according to equation 2, only easy accessible features such as areas and direction for a given floorplan are needed. Such information can be acquired from facility management services or observed from the outside of the building of interest.

The shown results in Table 4 build upon data from both Moroccan, German/ Swiss annotated examples. However, a classification based on distinct samples depending on the geographic areas improves the accuracy up to 88.01% and 87.43% for Moroccan and German/ Swiss data respectively. The feature weighting is varying accordingly. For the classification, an access to balcony
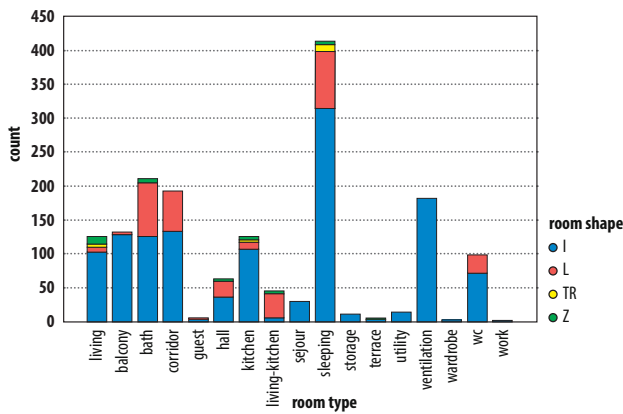
Figure 9. The distribution of functional use with regard to room shapes.

top features for room shape	weight
functional use	1
categorized area	0.91
proportional area	0.81
position	0.52
number of rooms	0.48
is corner room	0.39
direction	0.33
has balcony	0.29
....	...

Table 3. Excerpt of the most important features and their weights for the classification of room shapes (L and I).

is, for instance, for German/ Swiss area more relevant than for Moroccan housings. Hence, the variance of the induced weights is reflecting the influences of regional construction culture. A shape room classification based on examples from a specific functional use such as sleeping rooms, leads to more accurate results.

The classification results from this section and the previous section are of great interest in order to support and enrich the reasoning process for the derivation of indoor environments based on weak observations. In this context, the derived functional use can be applied to restrict the search space of housing room parameters. The search space can be narrowed based on statistical knowledge about location and shape parameters of rooms given their functional use. To this aim, the spatial relational database can be used to infer probability density functions. The latter can be approximated by Gaussian mixtures in order to enable probabilistic reasoning based on classical statistical algorithms. Hence, the combinatorial part consisting in, among others, finding corresponding windows to a-priori unknown rooms will take profit from the room functional uses. The search space can be further restricted based on the shape information of a given room. This is in particular of high relevance, if the geometric constraints observed from the outside are not strong enough to enforce some architectural patterns such as walls alignments. A further step worth of investigation is to perform a parametric learning of L-shaped room layouts. In this manner, statistical hypothesis tests can be performed in order to fit parametric probability density functions. The parameters of a probability density function which fits the shape parameters of specific room layouts well, can be used to represent different L-shapes.

Alternatively, other classifiers could be used for both classification tasks. Another possibility is to integrate the Bayesian classifier in a bagging or boosting process. In this context, a bagging-

	true I-shaped	true L-shaped	precision (%)
pred. I-shaped	235	54	81.31
pred. L-shaped	65	246	79.1
recall (%)	78.33	82	

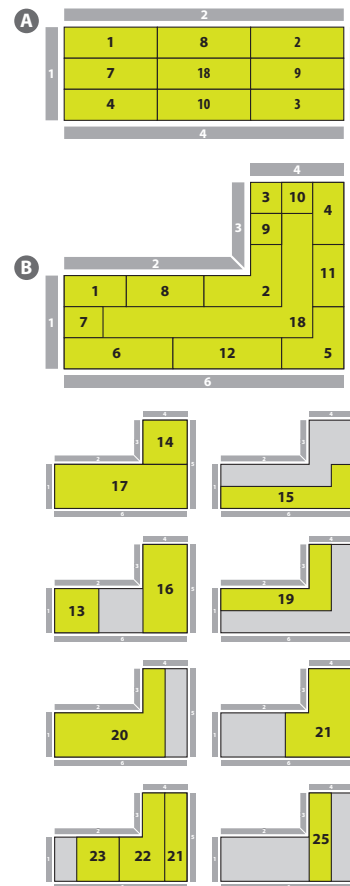Table 4. Classification results of room shapes. Two classes (I and L) are considered.

Figure 10. Possible encoded positions of a room within a floorplan. Nine possible positions in I-shaped floorplans (A). 25 positions are possible in L-shaped floorplans (B). The position encoding is used as a feature for the classification.

based approach (Breiman, 1996) with the Bayesian classifier delivered comparable results.

4   CONCLUSION AND OUTLOOK

This paper presented an approach for the prediction of room shapes and their functional uses. Based on sparse observations, we demonstrated that classification methods are suitable for the derivation of these two valuable findings. For the learning of a prediction model, openly accessible data has been used as training samples. Appropriate features have been extracted from an annotated spatial relational database. We took particular care to ensure that our features are easily to acquire from facility management services or to observe from the outside of the building. Sparse observations such as room areas and the direction of rooms turn out to be a good prior in order to infer room shapes as well as functional use. For the classification itself, we used a generative kernel-based Bayesian classifier which delivers probabilities of alternative class hypotheses. The classification results of the functional

use of housing rooms are extremely good, however, the classification of room shapes could be improved. To this aim, an analysis of the false positive predictions is a work in progress.

Our aim is to derive indoor models with as few observations as possible. This enables us to avoid costly capturing of high-resolution data. This reasoning approach consists of a combinatorial and a stochastic component. In contrast to office rooms, we believe that an extension of this approach with different room shapes and without room numbers is beneficial or even inevitable for housing rooms. In particular, an explicit addressing of uncertainty during the combinatorial part is of great interest. In this context, the probabilistic information derived by the classifiers are valuable. The incorporation of these results in the reasoning algorithm which derives the topology of floorplans for housing will be subject of an ongoing research.

ACKNOWLEDGEMENTS

REFERENCES

Becker, S., Peter, M. and Fritsch, D., 2015. Grammar-supported 3d indoor reconstruction from point clouds for "as-built" BIM. ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences II-3/W4, pp. 17–24.

Biljecki, F., Stoter, J., Ledoux, H., Zlatanova, S. and ltekin, A., 2015. Applications of 3d city models: State of the art review. ISPRS International Journal of Geo-Information 4(4), pp. 2842–2889.

Bishop, C., 2007. Pattern recognition and machine learning (information science and statistics), 1st edn. 2006. corr. 2nd printing edn. Springer, New York.

Breiman, L., 1996. Bagging predictors. Machine Learning 24(2), pp. 123–140.

Dehbi, Y., Loch-Dehbi, S. and Plümer, L., 2017. Parameter estimation and model selection for indoor environments based on sparse observations. ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences IV-2/W4, pp. 303–310.

Diakité, A. A. and Zlatanova, S., 2016. First experiments with the tango tablet for indoor scanning. ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences III-4, pp. 67–72.

Gröger, G. and Plümer, L., 2012. CityGML–interoperable semantic 3d city models. ISPRS Journal of Photogrammetry and Remote Sensing 71, pp. 12–33.

John, G. H. and Langley, P., 1995. Estimating continuous distributions in bayesian classifiers. In: Proceedings of the Eleventh conference on Uncertainty in artificial intelligence, Morgan Kaufmann Publishers Inc., pp. 338–345.

Khoshelham, K. and Díaz-Vilariño, L., 2014. 3D modelling of interior spaces: Learning the language of indoor architecture. ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XL-5, pp. 321–326.

Loch-Dehbi, S., Dehbi, Y. and Plümer, L., 2017. Estimation of 3d indoor models with constraint propagation and stochastic reasoning in the absence of indoor measurements. ISPRS International Journal of Geo-Information, 6, 90.

Mura, C., Mattausch, O., Villanueva, A. J., Gobbetti, E. and Pajarola, R., 2014. Automatic room detection and reconstruction in cluttered indoor environments with complex room layouts. Computers & Graphics 44, pp. 20–32.

Murphy, K. P., 2012. Machine Learning: A Probabilistic Perspective. The MIT Press.

Ochmann, S., Vock, R., Wessel, R. and Klein, R., 2016. Automatic reconstruction of parametric building models from indoor point clouds. Computers & Graphics 54, pp. 94–103.

Peter, M., 2017. Modelling of indoor environments using lindenmayer systems. ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XLII-2/W7, pp. 385–390.

Pintore, G., Garro, V., Ganovelli, F., Gobbetti, E. and Agus, M., 2016. Omnidirectional image capture on mobile devices for fast automatic generation of 2.5d indoor maps. In: 2016 IEEE Winter Conference on Applications of Computer Vision (WACV), pp. 1–9.

Prez, A., Larraaga, P. and Inza, I., 2009. Bayesian classifiers based on kernel density estimation: Flexible classifiers. International Journal of Approximate Reasoning 50(2), pp. 341 – 362. Special Section on The Imprecise Dirichlet Model and Special Section on Bayesian Robustness (Issues in Imprecise Probability).

Robnik-Šikonja, M. and Kononenko, I., 2003. Theoretical and empirical analysis of relieff and rrelieff. Mach. Learn. 53(1-2), pp. 23–69.

Rosser, J. F., Smith, G. and Morley, J. G., 2017. Data-driven estimation of building interior plans. International Journal of Geographical Information Science pp. 1–23.

Rosser, J., Morley, J. and Smith, G., 2015. Modelling of building interiors with mobile phone sensor data. ISPRS International Journal of Geo-Information 4(2), pp. 989–1012.

Scholkopf, B. and Smola, A. J., 2001. Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond. MIT Press, Cambridge, MA, USA.

Turner, E. and Zakhor, A., 2014. Floor plan generation and room labeling of indoor environments from laser range data. In: Computer Graphics Theory and Applications (GRAPP), 2014 International Conference on, IEEE, pp. 1–12.

Wand, M. P. and Jones, M. C., 1994. Kernel smoothing. Crc Press.

Wenming, W., Lubin, F., Ligang, L. and Peter, W., 2018. Miqp-based layout design for building interiors. Computer graphics forum.

Yue, K., Krishnamurti, R. and Grobler, F., 2012. Estimating the interior layout of buildings using a shape grammar to capture building style. Journal of Computing in Civil Engineering 26(1), pp. 113–130.

Zlatanova, S. and Isikdag, U., 2015. 3D indoor models and their applications. In: S. Shekhar, H. Xiong and X. Zhou (eds), Encyclopedia of GIS, Springer International Publishing, pp. 1–12.