Identification, Extraction and Three-Dimensional Building Model Reconstruction Though Faster R-CNN of Architectural Plans

Kunyang Ma[1,2(✉)], Yi Cheng[1,2], Wen Ge[1,2], Yao Zhao[2], and Zhang Qi[1]

[1] Information Engineering University, Zhengzhou 450000, China
394018925@qq.com

[2] Collaborative Innovation Center of Geo-Information Technology for Smart Central Plains, Zhengzhou 450000, China

Abstract. High-precision three-dimensional model is a high value information nowadays. Especially the 3D model of inner space of buildings can be used widely. However, it is not quite easy to acquire it. Building Information Model (BIM) provides a way to get detailed information of building to be build. But for most buildings already there the only thing that recording their internal structure we have are architectural plans. In order to get high precision model of building inner space, people scanning architectural plans into computer and vectoring every elements of building and reconstruction it by 3D building software as 3DMax and so on. Although this method is time consuming and high cost, it is the general method used in 3D building fields in recent years. In this study, we tried to speed up the modeling efficiency by identify and extraction information of building components automatically by deep learning algorithms. Faster R-CNN object-detection model, which was proved to be efficiency in image recognition is used in the paper to identity and extract independent building functional components such as tables, beds, cabinets and toilets and other equipment. This components are elements of plans but not exists actually in buildings. By deleting these "useless" components we can get a "clean" map of architecture and the vector data of building inner space can be obtained by image recognition algorithms as refine, extract center line and other operations. This paper proposed a method also to building the topology relationships of the vector data to drawn into room models in the buildings. And then, using room models to form building models. In order to verify the availability and generality of the model, we designed a method to transform the model we build to IFC, which is the famous BIM Standard model. Experiments shows that Faster R-CNN algorithm can provide high accuracy identification results, with the help of it the automation degree and efficiency of vectorization can be improved. The topology relationships defined to describe relations inside and between rooms are effective to form an entire building model. And the transform from 3D building model extracted from architectural plan in this way to IFC are feasible.

Keywords: Three-dimensional building model · Inner space · Faster R-CNN · Three-dimensional topology

Y. Xie et al. (Eds.): GSES 2019/GeoAI 2019, CCIS 1228, pp. 160–176, 2020.
https://doi.org/10.1007/978-981-15-6106-1_13

1 Introduction

Researches show that 90% of people's daily life is related to indoor space [1]. As the primary choice to express and reconstruct the indoor scene, three-dimensional model has the better user experience and the more realistic spatial analysis effect [2]. It is also the basis of realizing various indoor location services and the key integrating indoor and outdoor. At present, the mainstream methods of indoor three-dimensional modeling cannot meet the requirements of large-scale three-dimensional modeling rapidly [3–6]. Architectural plans can be used for indoor three-dimensional modeling, because of its low cost and rich content including geometry, size, material and the structure of building. However, many interference factors such as tables and beds are also in the plan. At the same time, most of the building models now available are ideographic [7], which lack the accurate description of spatial location and effective expression of the relationship between entities and is difficult to apply effectively to indoor location services and building management. Therefore, how to solve the problems mentioned above is the key issues of high-precision indoor three-dimensional modeling.

Guannan Li [8] and his group make up an integrated three-dimensional model in 3ds Max by MaxScript. They extract the information of layer objects, with the hierarchical relationship between them and main components retained, using the LISP command invoked from a secondary development interface of Tangent CAD.

After scanning and digitizing the architectural drawings, Yuanshu Li [9] builds an ideographic model with the walls extracted by using the sparse pixel vectorization algorithm, and building entities such as doors and windows which identified on the wall with the algorithm of LDA. The poor efficiency and accuracy of extraction and the lack of topological information in three-dimensional models make it difficult for the method in above studies, which extracting building structure information directly, to be widely used.

2 Methods

Based on the experience of previous studies and the idea of "reverse extraction", this paper proposes an efficient method which can identify and eliminate decorative components using object-detection technology, then establish the three-dimensional topological relationship among the entities in the building, which is the foundation for three-dimensional modeling and indoor location service. The diagram of main idea is shown in Fig. 1.

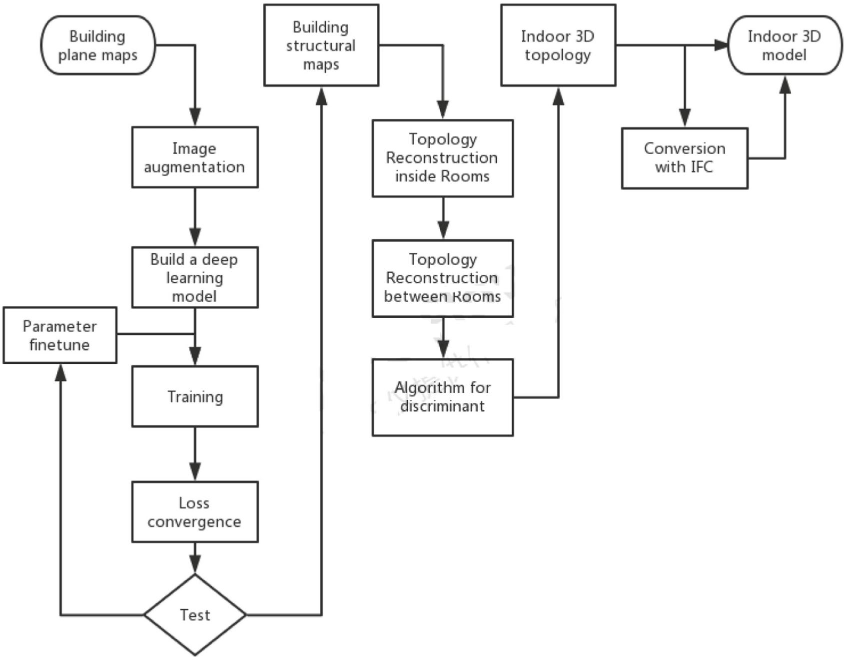

Fig. 1. Main idea of this paper

2.1 Geometric Extraction Based on Faster R-CNN

Characters of Architectural Plans

Compared with other data, architectural entities with unclear boundaries are dispersed in abstract forms in the architectural plan [10]. The abstraction, decentralization and overlap of information are important reasons for the difficulty of identification. Figure 2 shows the building plane map.

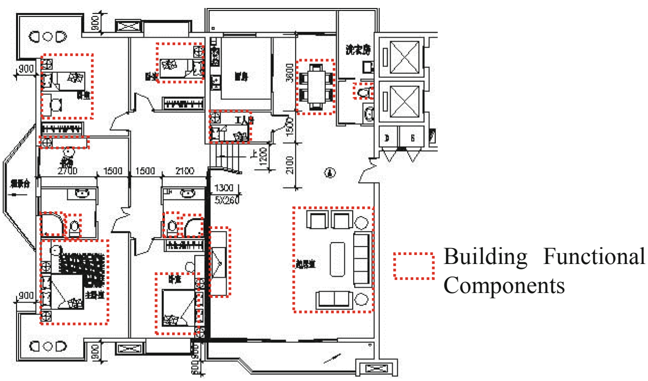

Fig. 2. Building plane map

Focusing on the research objectives of this study, building components in the architectural plan will be divided into two categories: architectural structural components and architectural functional components, according to the *Standard for Terminology of Civil Architectural Design* [11] and IFC (Industry Foundation Classes) standard.

Definition 1: Building structural components, which support the roof and floor and connect the rooms and floors in the building, includes walls, columns, doors, windows, railings and stairs.
Definition 2: Building functional components, which do not affect the internal structure of building, includes tables, beds, cabinets and toilets and other equipment.

Principle
Faster R-CNN model was proposed by Shaoqing Ren [12], where Region Proposal Network (RPN) is used to replace the Selective Search [13] method for rapid Region Proposal extraction. This model can predict what and where is the target through the training method of end-to-end. The anchor mechanism and border regression algorithm in the model make the recognition and prediction of mass data faster and more efficient.

(1) Anchor Mechanism
After receiving the feature map, a convolution layer with a convolution kernel of 3×3 will be used to convolute the feature map into feature vectors, at the same time, nine anchors according to three different areas and three different lengths are defined at the center of convolution core.
The feature vectors with anchors will be passed into a classified convolution layer with a convolution kernel of 1×1 to determine the scores of fore-ground (target) and background (non-target) in the anchors by Softmax classification function. Formulas are as follows:

$$s_i = \frac{e^i}{\sum_j e^j} \tag{1}$$

In this formula, i denotes the first element in a j-dimensional vector. As can be seen from the formula, the Softmax function maps a vector to a value on (0,1) and the sum is 1.

(2) Bounding-Box Regression
The feature vectors with anchors are simultaneously fed into a regression convolution layer with a convolution core of 1×1, simulating a fully connected layer, to predict the modified parameters including the displacement of the centers and the change of the edges of anchors, t_x, t_y, t_w, t_h. The predictive formulas are as given follows:

$$t_x = \frac{(G_x - P_x)}{P_w} \quad (2)$$

$$t_y = \frac{(G_y - P_y)}{P_h} \quad (3)$$

$$t_w = \log \frac{G_w}{P_w} \quad (4)$$

$$t_h = \log \frac{G_h}{P_h} \quad (5)$$

$$\mathrm{x} = w_a t_x + x_a \quad (6)$$

$$\mathrm{y} = h_a t_y + y \quad (7)$$

$$\mathrm{w} = w_a e^{t_w} \quad (8)$$

$$\mathrm{h} = h_a e^{t_h} \quad (9)$$

In the formulas, x and y indicate the coordinate of the center point of proposal region, w and h represent width and height of proposal region, G_i represents the true value, P_i represents the predicted value, $(i = x, y, w, h)$.

Procedures of Method

Based on the basic principle of Faster R-CNN, the main steps of geometric extraction model are shown as follows:

Step 1, image preprocessing. In order to remove the noise and highlight the main information of the architectural plan with different formats and wide sources, we removed the annotated axes and text descriptions first, then binary processing is done to obtain a black-and-white dot matrix.

Step 2, extract feature map. After the pretreatment, VGG-16 [14], a convolution neural network consisting of 13 convolution layers and 3 full connection layers are used as a feature extractor to generate the feature map for the follow-up Neural Networks. Figure 3 shows the structure of VGG-16 network.

Step 3, generate proposal regions. RPN, which plays a role in target location and image rough detection by anchor mechanism and bounding-box regression algorithm, is the most nuclear network of Faster R-CNN. When the feature map is transferred into the RPN network, the preliminary prediction of proposal regions (where and what are the targets) are obtained. The structure of RPN network is shown in Fig. 4.

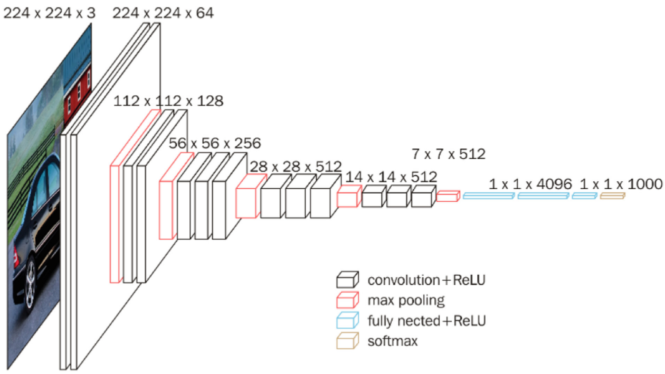

Fig. 3. The structure of VGG-16

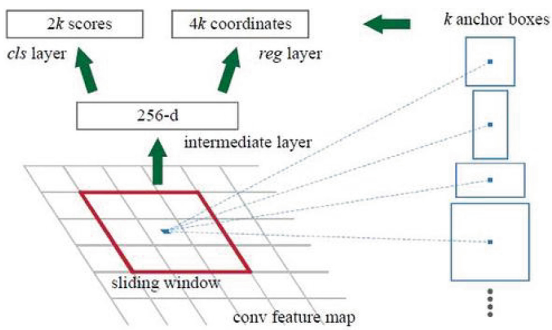

Fig. 4. The structure of RPN

Step 4, determine the target. The feature map in step1 and the proposal regions generated by RPN are simultaneously introduced into the network in this step to make secondary classification and regression. First, 300 target boxes are screened out from thousands of proposal regions by Non-Maximum Suppression (NMS) algorithm. Then, the target boxes are mapped to the feature maps and pooled into the same size and shape through the ROI pooling layer. Finally, the maps are transferred respectively into classification and regression network, which has the same structure and algorithm as RPN network, but the convolution layer is replaced by fully connected layer to abandon the weight sharing in different positions of convolution layer in order to realize further classification and better correction of target boxes.

Step 5, derive the results which contain the label and coordinate of targets that disturb indoor three-dimensional modeling from Faster R-CNN.

Step 6, remove the disturbances. A python script is used to automatically read the maps and delete the functional components by the graphical method according to the results obtained in the previous step to get structural maps of building internal.

The flow chart is designed as Fig. 5.

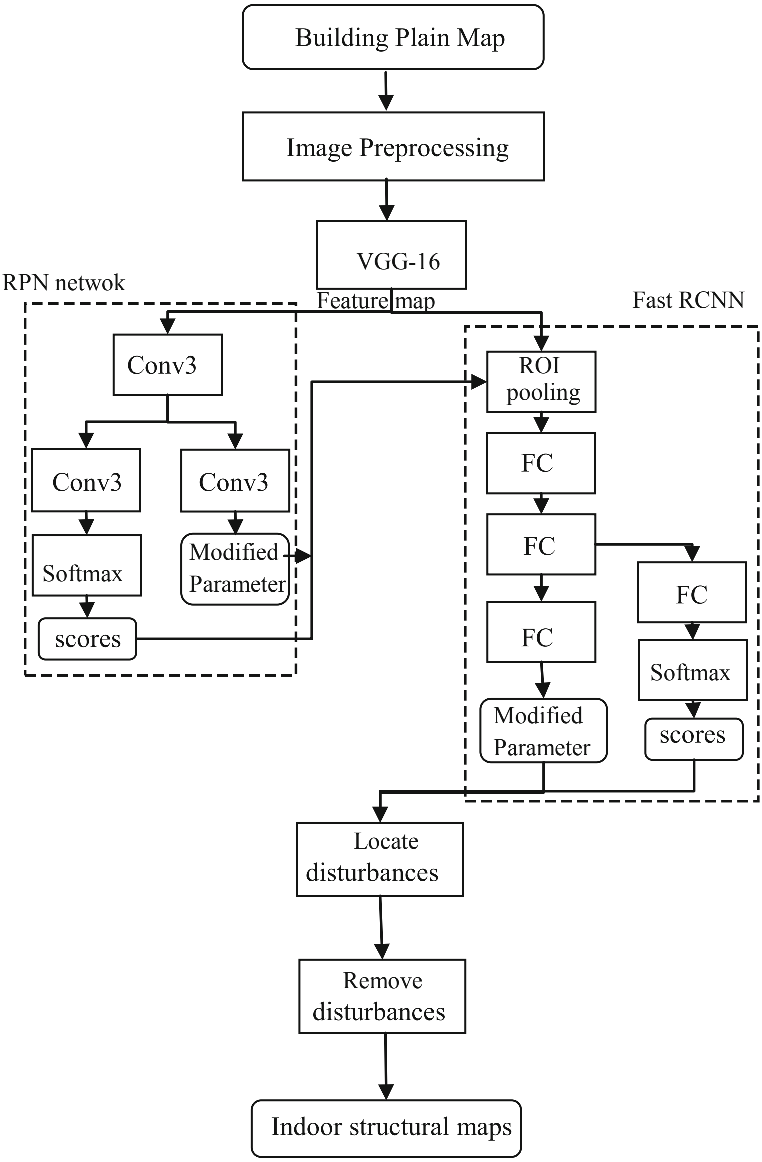

Fig. 5. The flow chart of geometric extraction

2.2 Building the Topological Model of Indoor Space

To build up a topological relationship of indoor space is essential for effective management of large-scale building [15]. Before that, we should transfer the ‘clean’ map obtained above to a vectorgraph by extracting center line. The topology reconstruction of indoor three-dimensional model is divided into two parts: inside rooms and between rooms. The topological from point to building in three-dimensional is established as shown in Fig. 6.

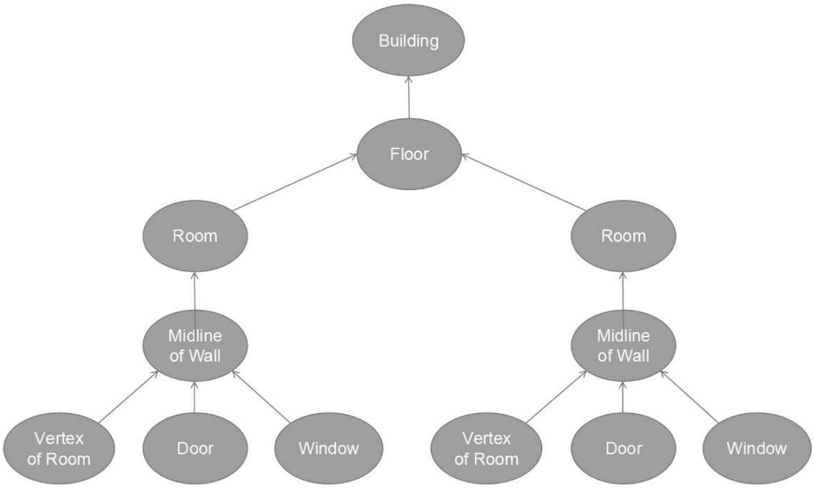

Fig. 6. Hierarchical structure of indoor topological model

Topology Reconstruction Inside Rooms

Rooms, corridors and stairwells that are considered as special rooms, are divided into vertex points, doors, windows, midline of walls and bottom of rooms, according to the different attribute types and functions. The attribute information of each element and the topological relationship between each other are established as Table 1 and Table 2.

Table 1. Attribute information of elements

Name	Attribute name
Vertex	Num
	Coordinate
	Floor
Door	Num
	Coordinate
	Width
	Height
Window	Num
	Coordinate
	Width
	Windowsill
	Top

Table 2. Relationships between elements

Relationship	Elements	Topological information
Component	Midline of walls and Vertexes	Num
		Start
		End
		Width
		Floor
	Midline of walls and bottom of rooms	Num
		Function
		Boundary
		Floor
Inclusion	Midline of walls and doors, windows	Num of midline of walls
		Num of doors
		Num of windows

Topology Reconstruction Between Rooms

Because there are many common parts in the boundary elements between rooms, which results in data redundancy and low efficiency of data management, it is necessary to establish three-dimensional topological relationship to uniquely duplicate elements. In the three-dimensional model of building interior, there are only two spatial relationships between rooms: adjacency and disjoint. In order to express the topological relationship between rooms more clearly, there are three kinds of situations in which the boundary elements exist in the common part.

(1) The two elements are completely common.
(2) The two elements are partially public, and there is a relationship of intersection or coverage between them.
(3) One element is completely contained in another.

According to the three categories above, spatial topological relationships between rooms can be classified into the following eight categories (Fig. 7).

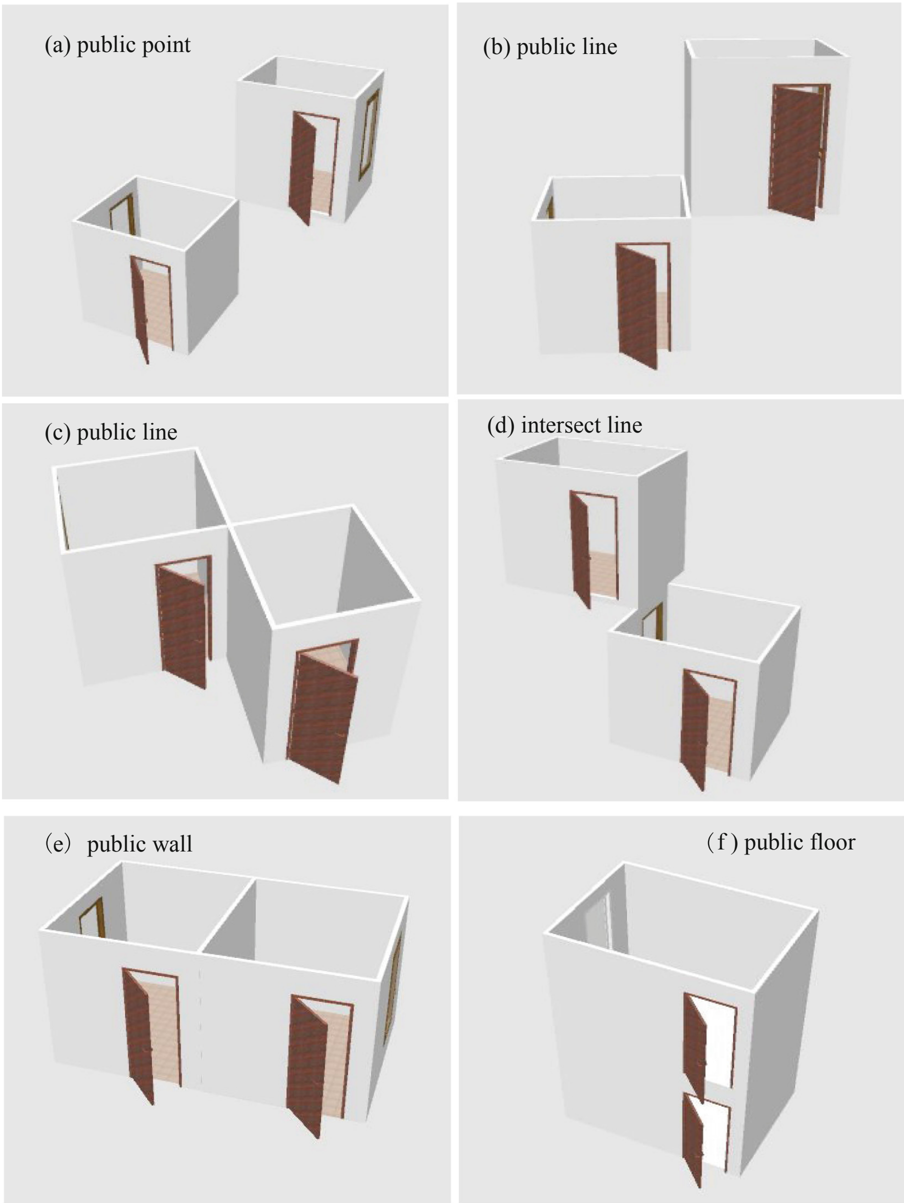

Fig. 7. Topological relationships between rooms

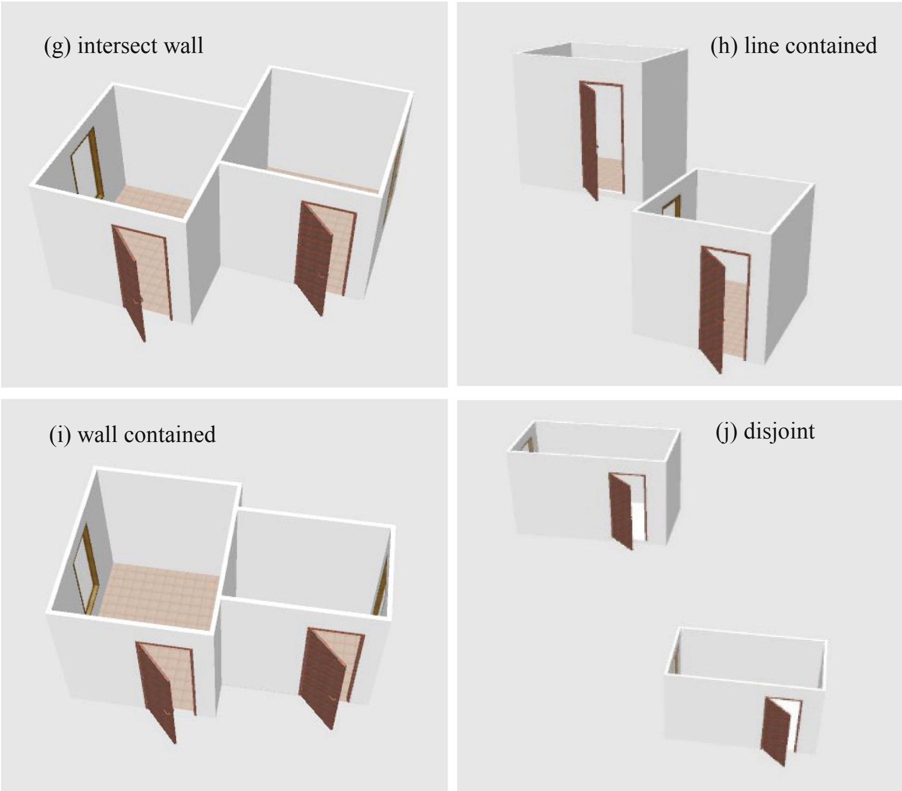

Fig. 7. (*continued*)

Based on the topological relationships has been showed above, an algorithm flow for judging is designed as Fig. 8:

After determining the adjacent relationship between two rooms, computational geometry is used to interrupt divide and delete the overlapping parts in the order from low to high dimensions, and attribute information is shared between the two rooms to improve the efficiency of data storage. In this process, new spatial elements will be generated, and the topological relationship inside the room needs to be updated to ensure the correctness and consistency of the data.

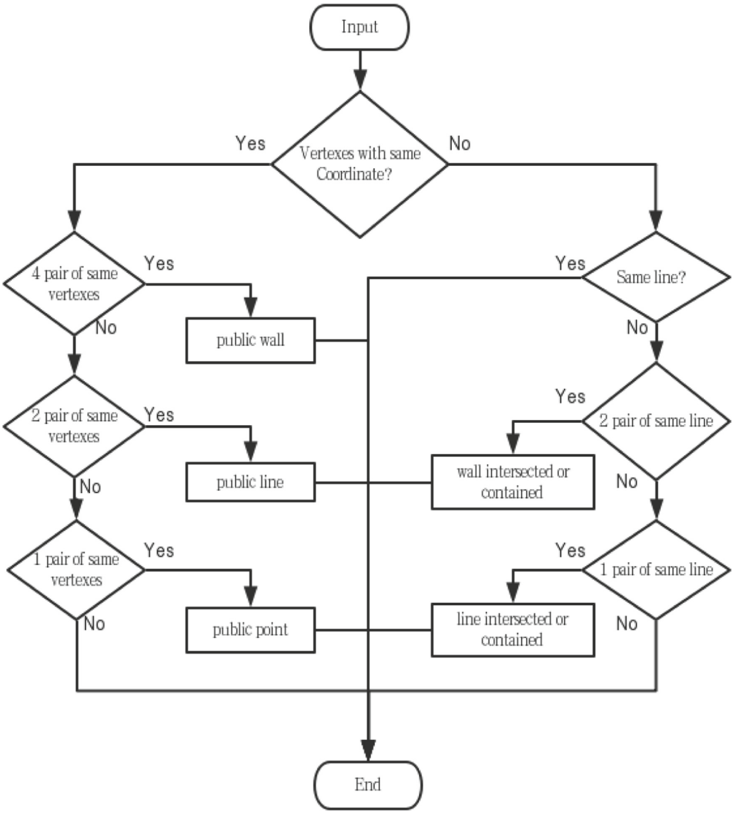

Fig. 8. The algorithm flow of judgment

2.3 Mapping to IFC

IFC (Industry Foundation Classes) which contains about 800 entities, 358 attribute sets and 121 data types in the latest version, defines a unified data format for building information to facilitate the interaction of data among applications in various industries, and it was first proposed by IAI (International Alliance for Interoperability) in 1997. Based on the indoor spatial topology model mentioned above, a transformation mapping between indoor spatial topology model and IFC standard (version 2×4) is established after analyzing the rules of how entity information and spatial relationship is described in IFC (see Table 3 and Table 4).

Table 3. Mapping rules of entities

Indoor spatial topology model	IFC
Vertex	IFCCartesianPoint
Door	IFCDoor
Window	IFCWindow
Midline of wall	IFCBoundedCurve
Bottom of room	IFCFaceOuterBound

Table 4. Mapping rules of relationships

Indoor spatial topology model	IFC
Vertexes to midline of wall	IFCRelAggregates
Doors to midline of wall	
Windows to midline of wall	
Midline of wall to bottom of room	
Public point	IFCRelConnects
Public line	
Public wall	
Components to room	IFCRelContainedInSpatialStructure
Rooms to floor	
Floors to building	

3 Experiment

3.1 Experimental Date

The data set contains 800 different building plans with no size requirement, each of which contains at least one object to be identified. The data sets are randomly divided into training sets used to fit models, cross-validation sets used to adjust model parameters and test sets for evaluating model performance in the ratio of 0.6:0.2:0.2. The experiment focused on the building component of beds and was carried out for better analyze the detection effect of the model, and data sets were divided into two categories: only objects and objects with interference in figure. Organization of data set is shown in Table 5.

Table 5. Organization of data set

	Only objects	Objects with interference	Total
Number of figures	314	486	800
Number of objects	345	1259	1603

3.2 Model Training

In this experiment, the parameters of the Faster R-CNN model are changed as follows: the number of candidate areas reserved for maximum suppression is set from 300 to 100 to improve training efficiency, the batch_size is set to 24 in order to take account of hardware computing ability and training efficiency, training times is set to 200,000. Training would be stopped when the model converges. The training set and verification set are used to train the model, and LabelImg is used to unify labeling of experimental objects on turning pictures firstly. After labeling, the training begins with about 1 s per step. 60 000 steps later, the model converges gradually and the loss value oscillates around 1. Trends of value are shown in Fig. 9.

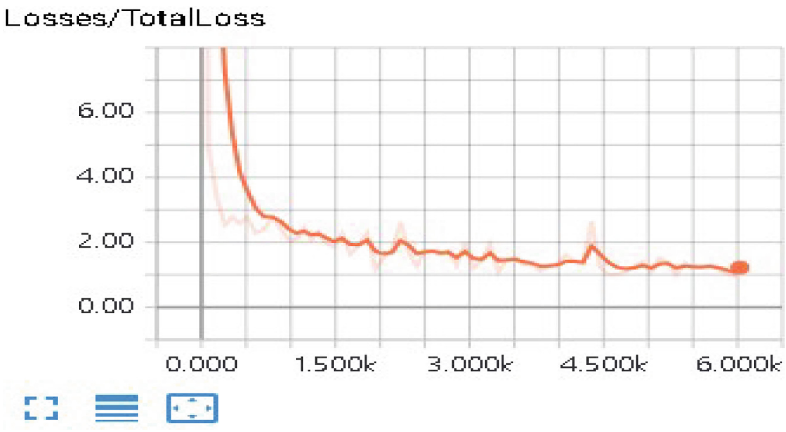

Fig. 9. Trends of value

After training, deriving the parameters of the model and testing the recognition effect of the model on test set. Precision means that how many of the detected items are right, recall means how many of the exact items are detected and mAP shows the comprehensive performance of model. The test results are shown in Table 6.

Table 6. Test results of model

Classification	Precision	Recall	mAP
Only objects	98%	91%	84%
Objects with interference	92%	83%	69%
Total	95%	87%	76%

3.3 Model Tests

After training, the plane maps of a two-storey building is used for testing. Firstly, the data is imported into the detection model to identify and delete the building's functional components, as shown in Fig. 10.

Fig. 10. Identify and delete building functional components

According to the method described in the third section, the indoor 3D topological relationship of buildings is built and converted to IFC standard file based on mapping rules. The contents of the document are shown in Fig. 11.

Fig. 11. IFC files

Finally, build up the three-dimensional model of building interior as Fig. 12.

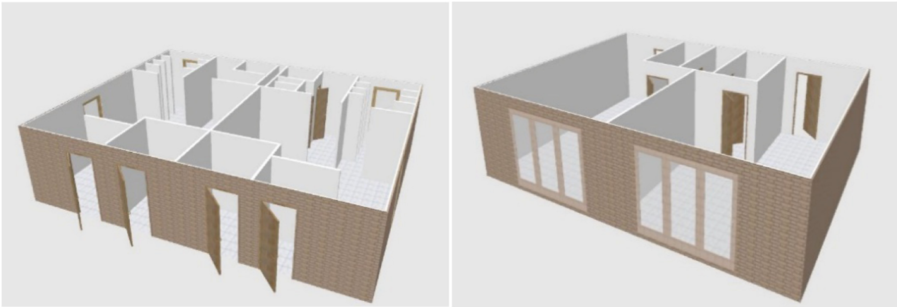

Fig. 12. Three-dimensional model of building interior

4 Summary

High-precision three-dimensional model is an important database for realizing integration of indoor and outdoor. On account of the idea of "reverse extraction method", this paper designs a detection model based on Faster R-CNN by using building plan for its accurate and cheap, which can automatically and efficiently obtain the internal structure map of building. At the same time, topology model of indoor space is established under the guidance of IFC, which can be used to accelerate three-dimensional modeling. Experiments show this method is suitable for three-dimensional reconstruction because of better accuracy and efficiency.

References

1. Worboys, M.: Modeling indoor space. In: ACM SIGSPATIAL International Workshop on Indoor Spatial Awareness, Chicago, p. 7 (2011)
2. Li, D.R., Liu, Q., Zhu, Q.: An integrated representation of outdoor and indoor in CyberCity GIS. Geomat. Inf. Sci. Wuhan University **28**, 253 (2003)
3. Isikdag, U., Zlatanova, S., Underwood, J.: A BIM-oriented model for supporting indoor navigation requirements. Comput. Environ. Urban Syst. **41**, 112 (2013)
4. Xue, M., Feng, L.I.: A framework for CAD/GIS/BIM online integration in project life-cycle management. Geogr. Geo-Inf. Sci. **31**(06), 30–34 (2015)
5. Cheng, L., Gong, J., Li, M., Liu, Y.: 3D building model reconstruction from multi-view aerial imagery and Lidar data. Acta Geodaetica Cartogr. Sin. **38**(06), 494–501 (2009)
6. Allen, P.K., Stamos, I., Troccoli, A., Smith, B.: 3D modeling of historic sites using range and image data. In: Proceedings of the IEEE International Conference on Robotics and Automation, ICRA, pp. 145–150. Taipei (2003)
7. Wang, X.F., Wang, Y.J.: Organization and scheduling of indoor three-dimensional geometric model based on spatial topological. Geomat. Inf. Sci. Wuhan Univ. **42**(01), 35–42 (2017)

8. Li, G.N., Wang, X.F., Zhang, C., et al.: Information extraction method for 3D building component based on T-arch. Geospat. Inf. **15**(07), 109–111 (2017)
9. Li, Y.S.: Fast 3D Building Modeling Based on Sparse Pixel Vectorization. Zhejiang University, Hangzhou (2013)
10. Hu, J., Yang, R.Y., Cao, Y., et al.: 3D reconstruction technology for architectural structure based on graphics understanding. J. Softw. (09), 1873–1880 (2002)
11. MOHURD. Standard for Terminology of Civil Architectural Design (GB/T50504-2009). China Planning Press (2009)
12. Ren, S., He, K., Girshick, R., Sun, J.: Faster R-CNN: towards real-time object detection with region proposal networks. IEEE Trans. Pattern Anal. Mach. Intell. **39**(6), 1137–1149 (2015)
13. Girshick, R., Donahue, J., Darrell, T., Malik, J.: Rich feature hierarchies for accurate object detection and semantic segmentation. In: IEEE Conference on Computer Vision & Pattern Recognition, p. 13 (2014)
14. Nagaoka, Y., Miyazaki, T., Sugaya, Y., Omachi, S.: Text detection by faster R-CNN with multiple region proposal networks. In: IAPR International Conference on Document Analysis and Recognition. ICDAR, Kyoto, p 15 (2018)
15. Zhu, Q., Li, D.R., Gong, J.Y., et al.: The design and implementation of CyberCity GIS. Geomat. Inf. Sci. Wuhan Univ. **01**, 8–11 (2001)