ICCEPM 2022
The 9th International Conference on Construction Engineering and Project Management
Jun. 20-23, 2022, Las Vegas, NV, USA

Indoor Semantic Data Dection and Indoor Spatial Data Update through Artificial Intelligence and Augmented Reality Technology

Sun Kwon[1]*

[1] *Department of Architecture, Technology and Enviornment, University of Illinois Urbana-Champaign, 117 Temple Hoyne Buell Hall, 611 E. Lorado Taft Drive, Champaign, IL 61820, USA,*
E-mail address: sunk2@illinois.edu

Abstract: Indoor POI data, an essential component of indoor spatial data, has attribute information of a specific place in the room and is the most critical information necessary for the user. Currently, indoor POI data is manually updated by direct investigation, which is expensive and time-consuming. Recently, research on updating POI using the attribute information of indoor photographs has been advanced to overcome these problems. However, the range of use, such as using only photographs with text information, is limited.

Therefore, in this study, and to improvement this, I proposed a new method to update indoor POI data using a smartphone camera. In the proposed method, the POI name is obtained by classifying the indoor scene's photograph into artificial intelligence technology CNN and matching the location criteria to indoor spatial data through AR technology.

As a result of creating and experimenting with a prototype application to evaluate the proposed method, it was possible to update POI data that reflects the real world with high accuracy. Therefore, the results of this study can be used as a complement or substitute for the existing methodologies that have been advanced mainly by direct research.

Key words: indoor spatial data, smartphone camera, semantics, points of interest, AR, CNN

1. INTRODUCTION

1.1. Research Background and Purpose

As indoor space has become taller and more extensive, interest in indoor building data, including indoor maps is grows. Among the components of indoor spatial data, indoor Point of Interest (POI) data should be updated in real-time according to the actual use of the space. The current method of acquiring due diligence is time-consuming and expensive, and it is challenging to keep the data up to date.

In this regard, I would like to suggest a way to keep the POI data in the indoor space updated using mobile data collected in real-time. With the spread of mobile devices, it is now possible to derive meaningful information from users' smartphone data being uploaded to the Internet in real-time. In addition, as the image classification technology using convolutional neural network (CNN), an artificial intelligence technology, has recently made significant progress, it is possible to identify the nature of space by classifying indoor scenes taken with a smartphone camera with CNN. Therefore, in this study, I classify indoor scene photos taken with a smartphone camera with

CNN to acquire meaningful information and devise a technique to match this information to indoor spatial data to update and build indoor POI data.

1.2. Related Studies

POI data is the core data constituting the spatial database and mainly refers to information on points of interest related to living convenience facilities (Ga, 2013). The indoor POI data to be covered in this study includes places, services, facilities, or events located tin specific locations in an indoor space and can better fully describe the indoor space (Claridades, 2019). In particular, when looking at its utilization, it is necessary to build and immediately update indoor POI data with enhanced semantic information for indoor navigation users (Ruta et al., 2015) that correspond to changes in information such as place use of indoor space, facilities, and services.

Recently, there have been attempts to build and update indoor and outdoor POIs in real-time by detecting space changes (Revaud et al., 2019) and classifying the nature of space (Ga, 2013) using photos taken of the space. Naver Labs (Revaud et al., 2019) showed the possibility of automatic updating of indoor POI through computer vision and artificial intelligence by detecting POI changes in signboards in the same space taken at different times at an indoor shopping mall. It has a limitation in that it is difficult to detect the semantics of many other types of spaces without text signs. This is the same as the limitation found in another previous study (Ga, 2013). Ga(2013) updated the outdoor POI by extracting the signboard text from the SLI (Street-Level Imagery) scene, so it is impossible to update the POI in a space where text is undetected with using method. Therefore, in this study, I tried to find a way to update the POI of various indoor spaces using spatial photos regardless of whether text is included in the “scene classification” technique.

In this study, to consider all kinds of indoor spaces regardless of whether text is included, actual space usage pattern is extracted as semantic information using scene classification of photos taken in the space, and the location criterion matched to the actual indoor POI data.

2. Study design

2.1. Definition and representation of indoor POI data

Indoor and outdoor POI data mainly refer to information on points of interest related to living convenience facilities (Ga, 2013), and indoor POI mainly include places, services, facilities, or events located in specific locations in indoor space (Claridades et al., 2019). In this study, attribute information related to the actual use of the indoor space (toilet, stairs, hallway, etc.) were limited to the POI name. There are two ways to represent the POI in indoor space data: representing a 3D shape of an object or space and representing a point at the center of the object or space.

In this study, indoor POI data were constructed and updated as point data with location information and indoor space usage patterns (semantic information) as POI names. Figure 2-1 is an example of expressing the indoor POI as a point in the indoor space data in the form of a vector map. One spatial polygon of the indoor spatial data geometrically includes one indoor POI point, and the two can be matched one-to-one.

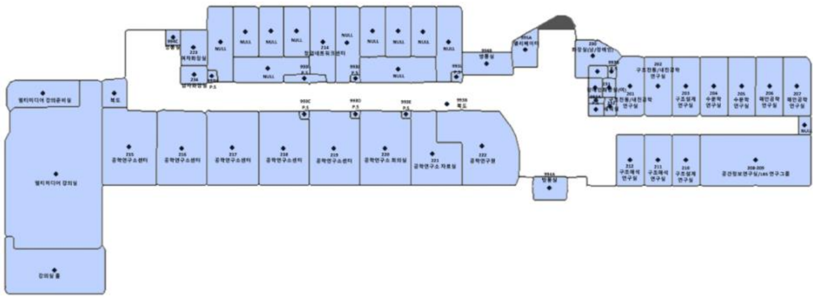


Figure 2-1. Example of indoor POI data in the form of points

2.2. POI matching technique between semantic information of indoor scene photos and indoor spatial data

Figure 2-2: Left shows how much spatial area is included in a picture of an indoor scene taken at a specific location (indicated by a green sphere). The results of Visual Access and Exposure (VAE) analysis, a type of Isovist analysis (Ostwald et al., 2018), is used to consider the average viewing angle of 70° of the smartphone camera, the viewing direction of the smartphone camera, and the actual wall acting as an obstacle in the field of view; The visible area of the smartphone camera is the same as the rectangular area shown in the figure. The visible rectangular area inside the circle expressed in Figure 2-2: Left represents the spatial area shown in the photo of Figure 2-2: Right.

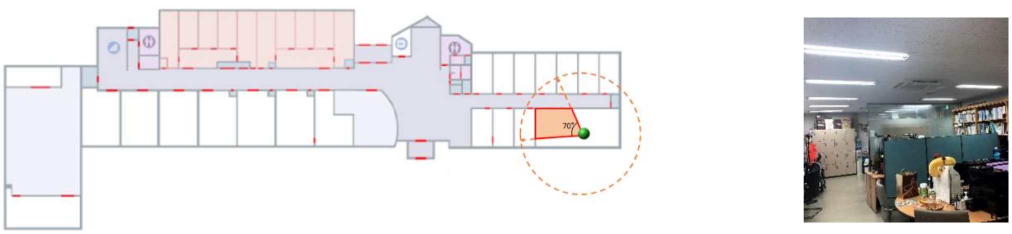


Figure 2-2. Left: 70 Isovist area in a closed space, Right: Photo corresponding to the area

The space shown in the photo is included in a larger space (polygon) surrounded by walls. Therefore, the attribute information of the POI data included in the indoor polygon data can be updated using the semantic information extracted from the photo. At this time, the semantic information is matched with the indoor spatial data using the location information of the previously obtained picture (Figure 2-3).

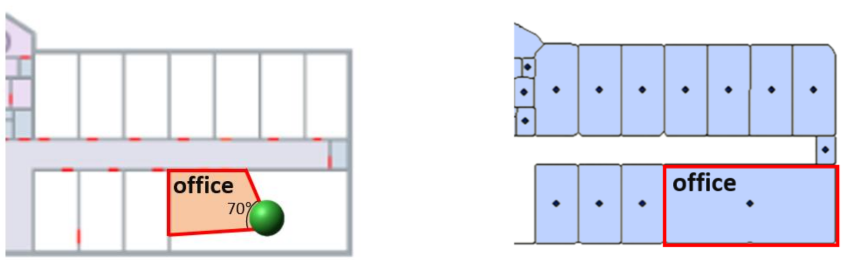


Figure 2-3. Left: The spatial area indicated by the photo, Right: Indoor polygon data including the space

For this, indoor polygon data, including the location point where the picture was taken, needs to be found, and this can be solved through a simple Polygon in Point (Sutherland et al., 1974) algorithm. In this study, the function was implemented using the ArcMap tool. Therefore, it is possible to update the POI of the corresponding area of the indoor space data using the photo acquired in the final closed space.

3. Experimental results and analysis

To evaluate the indoor POI data update method proposed in this study, an experiment was conducted using indoor space data and indoor scene photos of the 2nd floor of Seoul National University Building 35, the 1st floor of Building 65, and the 1st floor of Building 71. In Chapter 3, the analysis of actual experimental application and results and accuracy evaluation was performed.

3.1. Experimental Environment and Data

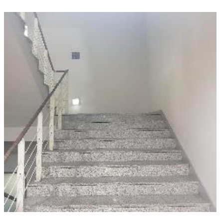
3.1.1 Experimental data

The indoor space data and the indoor scene photo were used for the experiment. The indoor space data is vector data in SHP format, and the room is expressed as polygons, and the POI is expressed as points so that they can be easily converted to indoor space data in other formats such as DWG, IndoorGML, IFC (Gelernter, J et al., 2019). This also has the advantage of building a large amount of indoor spatial data at a low cost.

3.1.2 Data Acquisition Scenarios

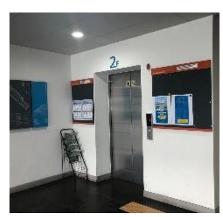
A low-cost indoor scene photo acquisition scenario is presented in this study. When a user walks toward a destination holding a smartphone vertically to use an AR-based indoor navigation application, the smartphone camera automatically takes pictures of indoor scenes at regular intervals and transmits them to the server. Then, the actual position coordinates (x, y, z) of the user (smartphone) can be obtained. Therefore, the indoor scene photos automatically taken at a specific point in time and at that location are transmitted and stored in the server, and the actual location coordinates information. In the actual experiment, a prototype application with only an indoor user location tracking function, AR-based indoor navigation, and a function to express on an indoor map was produced and used.

3.1.3 Extraction of semantic information from indoor scene photos using CNN



1.staircase(0.769)

2.bathroom(0.766)



3.elevator_lobby(0.300)

4.corridor(0.743)

5.staircase(0.520)

6. library/indoor(0.512)

7.office(0.313) 8. atrium/public(0.249) 9. office_cubicles(0.176)

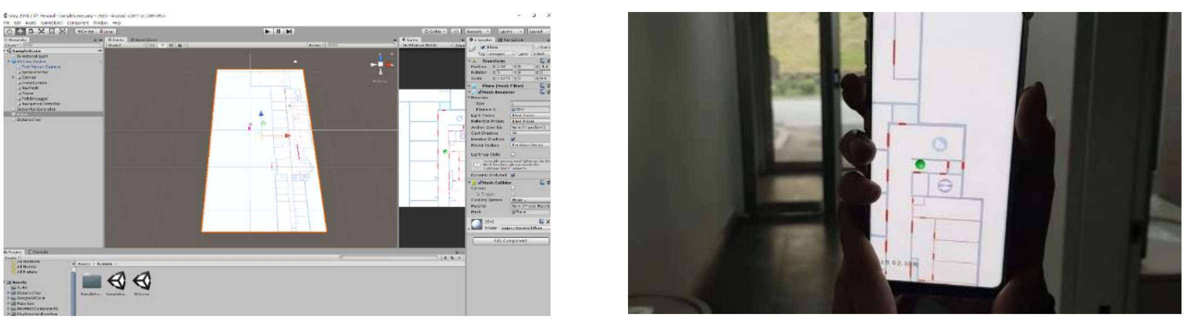
10.bathroom(0.804) 11.elevator_lobby(0.635) corridor(0.161) 12.home_office(0.268) office(0.212)

Figure 3-1. CNN Classification Results of Indoor Scene Photo of Seoul National University Building 35

In fact, Figure 3-1 shows an indoor scene photograph taken on the 2nd floor of the Seoul National University building 35, adjusted to 227 × 227 × 3 pixels to be used as input data of ResNet pre-trained with Places365 set, and then passed through a network. It is the result of classifying and extracting semantic information. The category name in which the photo was classified with the highest probability was defined as the semantic information of the matching space.

3.1.4 Estimating the location of indoor scenes using AR

In the data acquisition scenario presented above, indoor scenes are automatically captured while the user uses indoor navigation, and the user's location is estimated at the time of shooting and stored together with the indoor scenes. In this experiment, only the user's location estimation and expression functions were implemented to obtain the user's indoor location by making and using a prototype application, and an indoor scene was taken from the location and used as experimental data (Fig. 3-2).



Application creation screenshot Examples of using prototype applications

Figure 3-2. Prototype application production process and example of practical use

Using Google's ARCore, I created and used a prototype application to estimate the location of indoor scene photos and expressed the location on indoor space data. In this study, a Samsung Galaxy S8 (Android Version 9) was used, and an Android application was created using Unity3D (Version 2018.3.5.f1) in consideration of the convenience of development.

In the production process, the user's initial position and camera height (in this study, the user uses the smartphone a samsuming that indoor AR navigation was used, the camera height was set to 120 cm, which is the average eye line of sight height), a searchable space was set, and the set initial position and the position in the actual space were synchronized. The location of the user (device) was displayed as a green sphere on the indoor space data to increase visibility ([Figure 3-3]).

* Green sphere: indicates the location of the user (smartphone) estimated by mobile AR technology

Figure 3-3. When a user is in a general space surrounded by walls (bottom right)

Figure 3-4. shows the photos taken while moving to various places on the 2nd floor of Seoul National University Building 35, and the location is saved with the photos[2]. The photos' location information is displayed in relative coordinates with the upper left corner of the drawing as the origin. In this study, the direction and angle of the camera for matching with the indoor spatial data were manually inputted and used. However, ARCore and related smartphone sensors can recognize the orientation and rotation of the camera, so it is expected that automation will be possible in the process of commercialization in the future.

Figure 3-4. Acquisition of location information where pictures of indoor scenes were taken

3.2. Analysis of experimental results

The difference between the indoor scene photo of Seoul National University Building 35 Figure 3-5.Left on July 17, 2020, and the indoor scene photo taken on September 19, Figure 3-5.Right is noticeable. The space classified as office-cubicles and office on July 17 was classified as atrium/public in the photo on September 19.

Photos of interior scenes before POI update (2020.7.17) | Newly-taken interior photos (2020.9.19)

Figure 3-5. Changes in the use of space

Looking at the new indoor space data (Figure 3-6) where the POI of the existing indoor space data (Figure 2-1) has been updated, the POI name is “Concourse (atrium)” in “Building Manager Office.” You can check the corrected part with '.

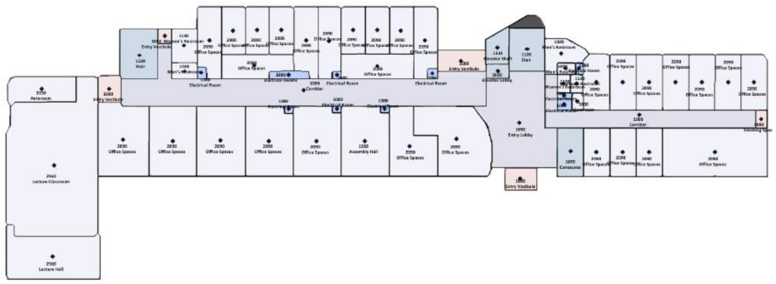


Figure 3-6. Indoor space data after POI update (2020.9.19)

3.3. Accuracy evaluation

The purpose of this study is to quickly and accurately build and update indoor POI data required in various spatial information service fields such as indoor map service and indoor navigation service in an efficient way. Therefore, it is necessary to compare the accuracy of the POI data generated by the method proposed in this study with the indoor POI data generated by the conventional method of acquiring due diligence. The indoor POI data constructed by the proposed method is evaluated through comparative analysis with the indoor POI data constructed through due diligence acquisition. The experimental area was the indoor space on the 2nd floor of Building 35, the 1st floor of Building 65, and the 1st floor of Building 71, Seoul National University.

As for the evaluation method, a cross-comparison analysis was performed between the POI data (Scene_POI) updated by the prototype system and the reference POI data (Reference_POI) updated by conducting a field survey in the same area (Ga, 2013). In order to verify how accurately the POI data constructed with the proposed method reproduces the spatial usage patterns in the real world, the evaluation index is based on recall and precision as shown in Equations (3-1) to (3-3). Furthermore, the F-measure was used (Do et al., 2003).

$$recall(\%) = \frac{Num(Reference\_POI \cap Scene\_POI)}{Num(Reference\_POI)} \times 100 \tag{3-1}$$

$$preision(\%) = \frac{Num(Reference\_POI \cap Scene\_POI)}{Num(Scene\_POI)} \times 100 \tag{3-2}$$

$$F-measure = 2 \times \frac{precision \times recall}{precision + recall} \tag{3-3}$$

Recall means the ratio of correctly extracted Scene_POIs to the number of Reference_POIs, and precision means the ratio of correctly extracted Scene_POIs among the extracted Scene_POIs.

Here, the actual Scene_POI data means a Scene_POI in which the POI name extracted from the indoor scene photo is matched to the corresponding polygon of the indoor space data. That is, Reference_POI means POI data with the exact location and POI name. Finally, the F-scale is a harmonic average that equally considers the importance of recall and precision, which has a reciprocal relationship, and has the meaning of overall accuracy.

<Table 3-1> shows the analysis result of Scene_POI for Reference_POI. Here, Scene_POI means POI data that has succeeded in automatically updating the POI name, and in this case, 85.06% and 86.05% in recall and precision, respectively, indicating an overall accuracy of 81.89%.

Table 3-1. Accuracy evaluation of indoor POI data of the proposed method

	Recall	**Precision**	**F-measure**
Scene_POI	85.06%	86.05%	85.55%

Table 3-2. An exampAccuracy evaluation of indoor POI data of the proposed method

Scene_POI		**Precision**	**F-measure**
True	**False**		
148	24	2	174

Looking at the results, the number of POI data existing in the target area, that is, the number of Reference_POIs was 174, but with the proposed technique, 172 POI data could be updated, and the number of accurately extracted and updated Scene_POIs was analyzed as 148 (<Table 3-2>).

4. Conclusion

In this study, to efficiently update and additionally build the properties of indoor POI data to be updated real-time, a practical methodology that can broadly consider various indoor space environments was proposed. Its validity was verified through experiments.

This study extracted semantic information close to the actual space usage pattern by adopting and using CNN scene classification technology to comprehensively consider the information contained in all kinds of indoor space photos taken with a smartphone camera. In addition, using cutting-edge AR technology, the location of the actual smartphone user was identified, and this was stored as location information of indoor scene photos. Using the information obtained from these two processes, the semantic information extracted from the indoor scene photos were matched with the existing indoor POI data based on the location, and their properties updated. Additionally, for this purpose, a scenario for acquiring low-cost indoor scene photo data based on smartphone users was presented, demonstrating the possibility of practical use.

In this process, new attempts were made to detect and divide the open space of indoor spatial data using spatial parsing and refining the indoor spatial data to flexibly reflect more diverse spatial usage patterns in the indoor spatial data.

As a result of the experiment, it was shown that the proposed technique could reproduce real world indoor POI data with an accuracy of 86.05% and an accuracy of 85.06%, indicating that the real world can be reflected with high accuracy. Therefore, this study presented a low-cost and efficient indoor POI data update methodology using the rich visual information and attribute information of indoor scene photos obtained with a smartphone camera. It has tremendous significance in that it presents a methodology closer to practical use.

REFERENCES

[1] Ga, Chillo. (2013). *A Method for POI Data Construction by Integration of SLI and Vector Maps* (Doctoral dissertation, Graduate School of Seoul National University).

[2] Claridades, A. R., Park, I., & Lee, J. (2019). Integrating IndoorGML and indoor POI data for navigation applications in indoor space. *Journal of the Korean Society of Surveying, Geodesy, Photogrammetry and Cartography*, *37*(5), 359-366.

[3] Do, H. H., Melnik, S., & Rahm, E. (2002, October). Comparison of schema matching evaluations. In *Net. ObjectDays: International Conference on Object-Oriented and Internet-Based Technologies, Concepts, and Applications for a Networked World* (pp. 221-237). Springer, Berlin, Heidelberg.

[4] Gelernter, J., & Maheshwari, N. (2019). Qualitative study of the incompatibility of indoor map file formats with location software applications. *Open Geospatial Data, Software and Standards*, *4*(1), 1-13.

[5] Ostwald, M. J., & Dawes, M. J. (2018). Isovist analysis, theories and methods. In *The Mathematics of the Modernist Villa* (pp. 95-124). Birkhäuser, Cham.

[6] Ruta, M., Scioscia, F., Ieva, S., De Filippis, D., & Di Sciascio, E. (2015, December). Indoor/outdoor mobile navigation via knowledge-based POI discovery in augmented reality. In *2015 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT)* (Vol. 3, pp. 26-30). IEEE.

[7] Revaud, J., Heo, M., Rezende, R. S., You, C., & Jeong, S. G. (2019). Did it change? learning to detect point-of-interest changes for proactive map updates. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 4086-4095).

[8] Sutherland, I. E., Sproull, R. F., & Schumacker, R. A. (1974). A characterization of ten hidden-surface algorithms. *ACM Computing Surveys (CSUR)*, *6*(1), 1-55.