Indoor-space Awareness using Descriptive data

Kang-jae Lee
Department of Geoinformatics
University of Seoul
Seoul, South Korea
kjlee@uos.ac.kr

Hye-Young Kang
Department of Geoinformatics
University of Seoul
Seoul, South Korea
hyezero@gmail.com

Jiyeong Lee
Department of Geoinformatics
University of Seoul
Seoul, South Korea
jlee@uos.ac.kr

Abstract— **Wide use of locational information has recently contributed to many GIS application services in a variety of ways, for example, shopping, tourism and transportation. In particular, these human behaviors have been widened from outdoor into indoor-space due to the appearance of large scale and complex buildings. Thus, indoor localization for various location based services has been researched in indoor-space using wireless LAN, RFID or Bluetooth instead of using GPS. However, these positioning technologies require intensive computation and have a limitation to retain stable accuracy. Thus, in this paper, we introduce Indoor-space Awareness Architecture using descriptive data such as names or phone numbers from images captured from mobile devices. OCR engine is especially applied in the architecture to recognize text from the descriptive data.**

Keywords - Indoor-space Awareness; Architecture; Descriptive data; Optical Character Recognition; Geocoding.

I. INTRODUCTION

The use of locational information has recently increased in web map services and map applications with a variety of information. As spatial information, the locational information is a basically used with non-spatial information for providing users with useful information such as shopping, tourism and transportation information near particular locations on the web map. Furthermore, as smart phones come into wide use, location-based applications are being developed steadily due to built-in GPS. So the locational information is used in real time received from GPS satellites. As a kind of positioning technology, GPS is the most useful system to obtain locational information for outdoor environments. However, in the case of indoor space, the GPS has a limitation on the accuracy of positioning since GPS signals cannot be directly come into buildings without any interference on the signals.

For the indoor positioning, several techniques are applied such as Wireless LAN [11, 12], RFID [4], Bluetooth [8] and InfraRed [3]. In spite of the application of these wireless techniques, awareness of location in indoor-space is restricted in the case of no infrastructure such as access point and sensors. Also, they have variable localization accuracy due to walls and columns. Thus, the alternative methods for the indoor localization need to be developed. In order to meet the minimum requirements, the alternative methods have to be simple to estimate locations and obtain locational data using information in indoor-space. So we designed Indoor-space Awareness Architecture using descriptive data which is readily available from stores, office rooms and convention halls in indoor space. So modified geocoding method is applied for matching descriptive data with locational data. In addition, for recognition and extraction of characters or words from the descriptive data, Optical Character Recognition (OCR) is applied. The OCR technology has grown with template matching and structure analysis in the pattern recognition field [13]. The descriptive data includes store names of complex shopping malls, room numbers of hotels or offices and phone numbers for deriving locational information in indoor-space.

Implementation of this research is conducted on Android OS for smart phones. Smart phones are widely used and most of them have camera modules which are appropriate for capturing images simply and quickly for acquiring descriptive data. As a result, an application based on Android OS is developed according to the structure of Indoor-space Awareness Architecture. Furthermore, a query service based on the architecture is implemented using related search terms.

II. RELATED WORKS

Numerous researches have been carried out regarding location awareness and navigation using a variety of data and technologies. For this paper, the main issue is how image data is handled with locational information using geocoding, thus the previous researches on image-based positioning and geocoding methods are reviewed.

Image-based location awareness has been developed with geo-tagged image data. The research on the algorithm for estimating geographic information from image data was conducted using a purely data-driven scene matching approach [9]. According to the research, over 6 million GPS-tagged image data was stored in database and some of them were partially used for the evaluation of the method. Images for the test were matched with database in accordance with color patterns and texture features so as to recognize correlation of between the images and geographic locations. And the geo-locations of the images were estimated using a set of GPS tagged nearest neighbors. Another research on the system for managing digital photographs with geographic location tags was carried out [14]. The digital photographs with locational information include semantic content which helps searching and indexing. Furthermore, the research proposed many kinds of methods for acquiring location metadata. One of them is using digital cameras or mobile

phones with built-in GPS chips. So latitude and longitude information is stored in the metadata of image. Another method is to obtain locational information from descriptive text. Most of landscape or scenery image data embedded in web pages generally expresses the locations, buildings or landmarks with text data. Thus, the descriptive text data is used as geographic location tags for the images.

Generally, geocoding is the process of deriving geographic coordinates from geographic data such as addresses or postal codes matching it with reference database. The geocoding goes through three processes which are parsing, matching and locating[10]. Address data is converted to standardized format in the parsing process. And next the input data is matched with pre-built reference database and result records are selected in the matching process. Finally, in the locating process, geographic coordinates of the locations are extracted from the result records. In the processes, various geocoding methods are applied and in terms of the methods, the address matching technique[7], and the 2D and 3D area-based address matching technique[10] have been researched.

In this research, first of all, we concentrated on the indoor environments GPS signal cannot be received accurately. So, instead of using GPS, we chose to use descriptive data having implicit geographic locations. In other words, the descriptive data can be used to derive explicit geographic coordinates in a building. Thus, secondly, during this process, additional process is required for the extraction of descriptive data from images captured by users. So we decided to use OCR engine for extracting words and numbers from store names, room numbers and phone numbers in convention halls and complex malls.

III. SYSTEM ARCHITECTURE

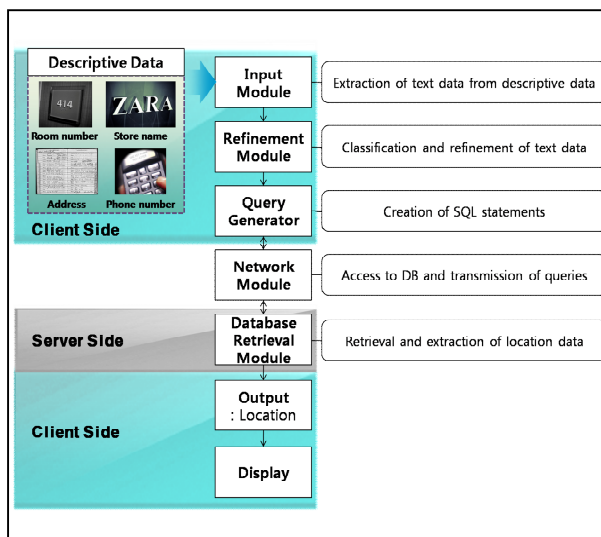In this paper, Indoor-space Awareness Architecture is designed as shown in Figure 1.

Figure 1.   Indoor-space Awareness Architecture.

Among the whole processes, first, the input module extracts text data from descriptive data. Second, the text data is classified into several categories such as an phone number, room number and store name through the refinement module. The text data is parsed to meet the standardized form in order to match it with reference database. And then, query statements are created by the query generator for searching corresponding attributes in database. The query statements are transmitted to the network module, and next retrieval and extraction of location data using the queries are performed in the database retrieval module. The output data is finally displayed in the client side in recognizable form such as polygon or point on a digital map.

A. Descriptive Data Input and Refinement

As a first component shown in left side of Figure 2, the input module receives descriptive data such as store names, phone numbers or room numbers in complex shopping malls or convention halls.
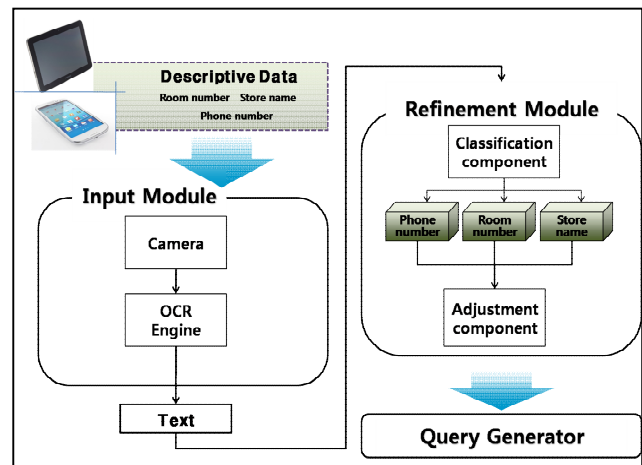

Figure 2.   Input and Refinement module of Indoor-space Awareness Architecture.

The camera module captures image data from camera-equipped mobile devices, and then compresses and stores it as image formats. And next, the image data representing descriptive data is sent to the next stage, OCR engine. The OCR engine extracts text data from the image data through several complex processes. OCR is generally divided into on-line and off-line character recognition depending on data acquisition methods. Among them, off-line character recognition is the method that text data is acquired from the fixed static shape of the character in printed documents and signs using a scanner or camera. So it is applied to this research for acquiring text data from images, which represent fixed static shape of characters. Off-line character recognition generally consists of four steps. Firstly, pre-processing adjusts the slant of documents and explicitly divides characters from background for increasing recognition accuracy. Secondly, segmentation is a process to split character strings into many rows, words and characters. Thirdly, a process of feature extraction analyzes the pattern of each character using various methods and then characters

are individually recognized. Lastly, post-processing increases recognition accuracy taking into account semantic information of context. Incorrectly recognized words are proofread using dictionary data method[5] or HMM method[6]. After the processes of the input module, as a result, the text data is extracted.

As a second stage in the Indoor-space Awareness Architecture, text data is classified into a phone number, room number or store name in the classification component as shown in the right side of Figure 2. The reason for necessity of classification is that it allows the text data to be refined correctly depending on the content types of individual text data. For example, phone numbers consists of a combination of numbers and it is at least greater than minimum digit numbers including area codes depending on many countries without any letters. Afterwards, classified text data is adjusted for the standardized structure. And then, queries are created by the query generator for searching corresponding attributes in database.

B.   Searching and Display

Generated queries are transmitted from client side to server side through the network module as shown in Figure 4. The network module connects the client side to the reference database of the server side. It accesses the databases and sends the queries to the database retrieval module. And then, the database retrieval module searches and extracts location data using the queries. In this module, the query manager checks validation of the queries and decides how to access reference database. Next, it is decided that which kind of format the location data extracted from reference DB has in the formatting component. The output data is finally displayed in the client side in recognizable form such as a polygon or point on a 3D digital map.
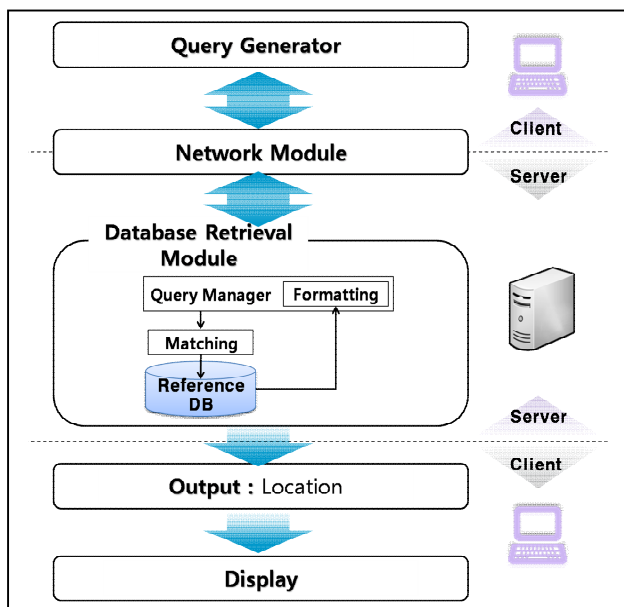

Figure 3.   Database Retrieval module and output data of Indoor-space Awareness Architecture.

The reference database schema is designed as shown in Figure 4. The 'Feature' table contains a set of attributes such as an identifier (ID), floor number, room number, coordinates and foreign key(R_ID) which refers to the identifier(ROOM_ID) of 'Room' table. 'Room' table contains a set of attributes such as a phone number, additional number for the same name of stores(ST_ORDER) and foreign key for referring to the identifier of 'RoomName' table. 'RoomName' table is designed including attributes of store names with various languages respectively.
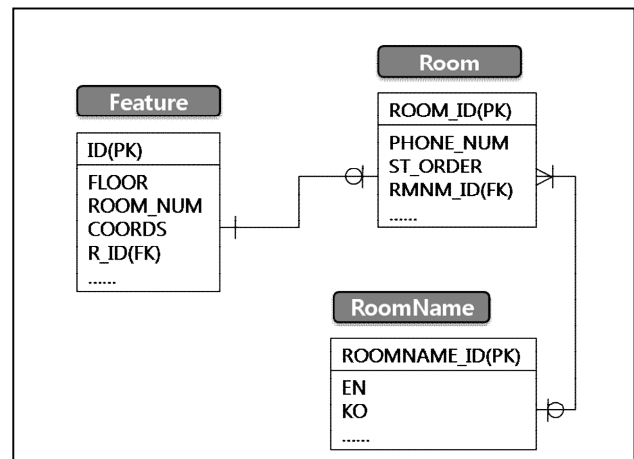

Figure 4.   Schema of Reference DB.

Especially, matching process of store names works according to the algorithm below. Firstly, a store name *snt* is directly matched with the list of corresponding attribute. If it is matched at the first attempt, (*x*, *y*, *z*) coordinates are extracted without any additional works. However, if it cannot be matched perfectly, it is partially matched and the location data linked to *output* is extracted. Regardless of the partial matching, if the *output* has candidates or a matched attribute data cannot be found, then matching process operates manually.

```
Input: Boolean b;
       Array output;
Procedure MatchingProcess(Storename snt){
   output←CALL FullMatching(snt)
   IF (output.length = 1 )
      CALL Location(output)
   ELSE IF (output.length = 0 )
      FOR each attribute att from ReferenceDB
         b←CALL IsContain(snt, att)
         IF (b)
            output←att
         END IF
      END FOR
      IF (output.length = 1)
         CALL Location(output)
      ELSE
         CALL ManualMatching(snt)
      END IF-ELSE
   END IF-ELSE
}
```

IV. IMPLEMENTATION

Our research is executed based on the Indoor-space Awareness Architecture. A building for the research spot is COEX, which has a largest underground mall and convention centers on the ground. For the implementation of the architecture, an application program is developed based on the android app using Open source OCR program called 'Tesseract'[2][1]. For the actual implementation of the research, as a first stage, sample data is stored in the reference database. As shown in Figure 5 (a), the name of store can be captured as text data using equipped camera in a smart phone. And next, the text data is automatically matched with the store name from the reference database. Then, the current position is represented as a polygon in 3D visualization with information about the store. The keypad enabling manual matching is supported with auto complete text view as shown in (b). For the application of our research, we implemented a query service and the result areas are represented as distinctively colored polygons like (c).

(b) manual matching

(a) text matching and 3D visualization
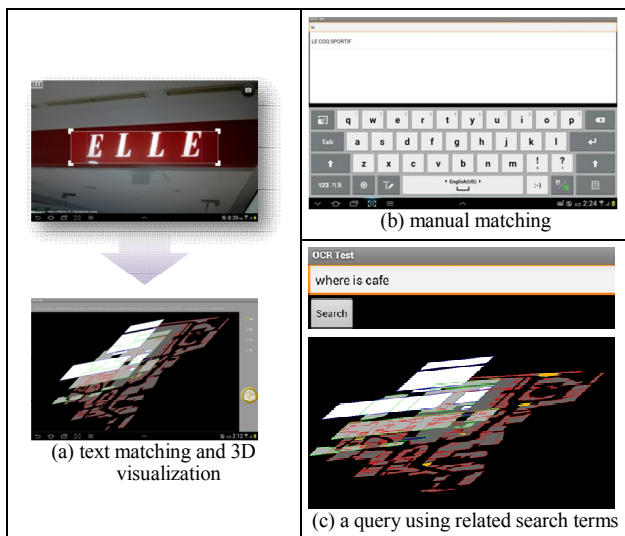
(c) a query using related search terms

Figure 5. Indoor-space Awareness with 3D Visualization.

Correctly recognizable rate of locations using all kinds of descriptive data is 94.35%. Specifically, the automatic matching rate of English storefront signs is 82.41% using 108 samples, whereas such matching rate of phone numbers is 100% using 10 samples. Although the number of samples is too low in the case of phone numbers, it is analyzed that the error rate in the case of the storefront signs is caused by illumination, various fonts and background color.

V. CONCLUSION

This research introduced Indoor-space Awareness Architecture using descriptive data for detecting and extracting locational information through modified geocoding method. Among the process, OCR engine is especially applied to recognize text data from the descriptive data. For further researches, special processes should be researched and used in the OCR program for increasing recognition rate. For example, adaptive thresholding needs to be applied for eliminating the effect resulting from illumination. And, for various application of the architecture, the research on incorporating 3D network model can be possible.

REFERENCES

[1] Android-OCR. https://github.com/rmtheis/android-ocr

[2] Tesseract open source OCR engine. http://code.google.com/p/tesseract-ocr.

[3] Baus, J., Kruger, A., Wahlster, W. 2002. A Resource-Adaptive Mobile Navigation System. 2002. International Conference on Intelligent User Interfaces IUI02.

[4] Bekkali, A., Sanson, H., and Matsumoto, M. 2007. RFID indoor positioning based on probabilistic RFID map and kalman filtering. In Proceedings of the IEEE International Conference on Wireless and Mobile Computing, Networking and Communication, 21-27.

[5] Bokser, M., 1992. Omnifont technologies. Proc. IEEE, vol. 80, pp. 1066-1078.

[6] Chen, M. Y., Kundu A., and Zhou J. 1994. Off-line handwritten word recognition using a hidden Markov model type stochastic network. IEE Trans. Pattern Anal. Machine Intell., vol. 16, pp. 481-496.

[7] Drummond, W.J., 1995. Address matching: GIS technology for mapping human activity patterns. Journal of the American Planning Association, vol. 61, pp. 240-251.

[8] Feldmann, S., Kyamakya, K., Zapater, A., and Lue, Z. 2003. An Indoor Bluetooth-Based Positioning System: Concept, Implementation and Experimental Evaluation. In Proceedings of the International Conference on Wireless Networks, 109-113.

[9] Hays, J., and Efros, A. A. 2008. Im2gps: estimating geographic information from a single image. In Proceedings of the IEEE Conference Computer Vision and Pattern Recognition.

[10] Jiyeong Lee, 2009. GIS-based geocoding methods for area-based addresses and 3D addresses in urban areas. Environment and Planning B : Planning and Design, vol. 36, pp. 86-106.

[11] Kaemarungsi, K., and Krishnamurthy, P. 2004. Modeling of Indoor Positioning Systems Based on Location Fingerprinting. In Proceedings of the IEEE INFOCOM, vol. 2, 1012-1022(May 2004).

[12] Li, B., Salter, J., Dempster A. G., and Rizos, C. 2006. Indoor positioning techniques based on wireless LAN. In 1st IEEE Int. Conf. on Wireless Broadband & Ultra Wideband Communications, Sydney, AUS(Mar. 2006).

[13] Mori, S., Suen, C. Y., and Yamamoto, K. Sannella, M. J. 1992. Historical Review of OCR Research and Development. In Proceedings of the IEEE, vol. 80, no. 7, 1029-1058.

[14] Toyama, K., Logan, R., and Roseway, A. 2003. Geographic Location Tags on Digital Images. In Proceedings of the International Conference on Multimedia, 156-166.