Managing and Retrieving Spatial Information in Architectural Floor Plan Databases

Jie-Eun Hwang¹ and Jin-Won Choi^{*2}

¹ M.S., Department of Housing and Interior Design, Yonsei University, Korea

² Assistant Professor, Department of Housing and Interior Design, Yonsei University, Korea

Abstract

This research investigates spatial information retrieval in architecture focused on constructing efficient metadata that is crucial for data retrieval. Generally speaking, metadata is 'structured data about data' to describe resources especially in digital realm. In this research, metadata is a sort of data object useful in searching spatial information, to describe raw spatial data objects with not only attribute but also content consisted of structure and semantics. There are two issues that motivate this research; 1) how to represent intangible space as a form of data object, and 2) how to search and browse these data objects in a huge database, an issue of content-based information retrieval. We analyze spatial information of a floor plan, especially focused on the apartment unit floor plan common in Korea. Then we extract the metadata items in a structured manner. To manage the items efficiently, we develop a data model for spatial information according to the concept of "Structured Floor Plan." For exploiting the metadata, this research shows several possibilities of query operations to present a set of sample queries about L-D-K(Living room - Dining room - Kitchen). Implementation of the prototype system is divided into three parts: 1) a modeling module, Vitruvius; 2) an indexing module, SpaceCore; and 3) a browsing module, SpaceScope.

Keywords: content-based retrieval; metadata; structured floor plans; RDF; spatial information

1. Introduction

1.1 Backgrounds and Purpose

The rapid development of information technology has incited the area of information retrieval (IR for short) to grow rapidly. These days, the scope of IR has expanded far beyond its original purpose, such as indexing and searching for useful documents. IR includes classification and categorization, systems architecture, user interfaces, data visualization, filtering, languages, etc. Since the World Wide Web was introduced, it has become a universal repository of human knowledge. The huge amount and various types(e.g. multimedia) of information lead IR to new challenges such as the heterogeneity of data, the fuzziness of information, the loss of information in the creation of indexes, and the need for an interactive refinement of the query result (Baeza-Yates and Ribeiro-Neto 1999).

IR is also a core issue in the architectural domain. Regards to the architectural information, research has been done actively within the digital paradigm such as STEP, CALS, and IFC. However, by now, the research tends to be inclined to how to convert conventional data

into digital form efficiently without any loss or how to standardize such data. Concerning IR, research is merely confined within the level of data retrieval. Moreover, although space is fundamental in architecture, the majority of research mentions only data about surroundings of space, such as building components, drawings, specifications, equipments, etc. (Kim 1996; Zhou et al. 2002). There has been very limited research on spatial information.

There are two issues that motivate this research; one is the content-based information retrieval, another is spatial information as content for retrieval. Although knowledge of the spatial content of a building is most important in architecture, there has been no logical method to control and manage it.

This research investigates the spatial IR in architecture focused on two topics; 1) constructing an efficient metadata that is crucial for data retrieval and 2) managing the metadata with knowledge-based query operations in the viewpoint of the content-based IR. Metadata is 'structured data about data' to describe the resources especially in digital realm (Gill 2000). Especially it is a sort of information on the organization of the data, the various data domains, and the relationship between them (Baeza-Yates and Ribeiro-Neto 1999). Demands for metadata grow rapidly, especially in such an area of multimedia IR because of the needs of the different querying paradigm, the adequate and efficient processing techniques of retrieval, and the semantics of multimedia

*Contact Author: Jin-Won Choi, Assistant Professor, Dept. of Housing and Interior Design, Yonsei University, 134 Shinchon-dong, Seodaemun-gu, Seoul, Korea
Tel: +82-2-2123-3137 Fax: +82-2-313-3139
e-mail: jchoi@yonsei.ac.kr

(Received November 8, 2002 ; accepted March 15, 2003)

data (Boll et al. 1998). Similar to the recent multimedia information retrieval approaches, this spatial IR system should be able to afford users to not only data attributes but also object's content.

2. Methodology and Scope of Process

This research refers to the spatial content of a floor plan as that of a building or space. Conventionally, the floor plan is typically used to represent spatial information. In spite of the limitations of two-dimensional media, a floor plan is well organized with a standardized format of symbol and carries so many meanings, from an intuitive concept to a detail indication. Much spatial information can be extracted from the floor plan. Therefore, this research analyzes the content of space with the floor plan and derives metadata of spatial information from it. As a sample, an apartment unit floor plan is chosen, a common residential space widely spread in Korea. The reason to be chosen is that it has a rational size and composition of spaces to be investigated in this research, also the demand for retrieval is higher than with other types of buildings.

First, the researcher analyzes the spatial information of an apartment unit floor plan and classifies it to construct the data model for metadata. And then the study goes on in order to investigate how to generate metadata from that analysis, how to organize metadata, and how to exploit metadata. To organize and describe metadata, the RDF is adopted.

The prototype spatial IR system is implemented on the web, and Microsoft SQL Server 2000a is used as database with its XML applications. As a metadata generator, Vitruvius is used to produce structured floor plans.

3. Classifying Spatial Information

3.1 The Needs for the Metadata

Metadata is 'structured data about data' which describes resources, especially in the digital realm (Gill 2000). In particular, it is a type of information about the organization of data, the various data domains, and the relationship between them (Baeza-Yates and Ribeiro-Neto 1999).

The demand for metadata is created by the specific need for spatial information characteristics beyond the capacity of a traditional role of retrieval. The metadata of multimedia is a good paragon for that of spatial information. Metadata plays a far more important role in managing multimedia data than does the management of traditional well-structured data or information retrieval techniques applied to text-only data. This is due to the different query paradigm, inadequate processing techniques, lacking efficiency, and semantics of spatial data (Boll et al. 1998). In another words, the exact-match paradigm for query is no longer suitable. Various types and huge amounts of digital media increase the need for new processing such as content-based retrieval. It has consequences on effective management and retrieval of

digital media. Semantics of multimedia data like images, video, and audio is implicit to the raw media data. Analyzing and processing such data, allow semantics to be made explicitly to some extent on a different abstract level, from feature values to knowledge-based concepts. Metadata describing this semantics explicitly may still not be sufficient for exact-match querying. For example, feature values of color distribution in an image for red, black, and yellow still do not allow the conclusion that the image depicts a sunset. A different query paradigm is needed to allow a proper mapping of the user's ideas to the available semantics of a media object and, finally, to the raw data.

The new needs and challenges for metadata of spatial IR in architecture are similar to those of multimedia. Because the media for spatial data, such as floor plans drawings, 3-D models, and rendered images, are technically similar to the case mentioned above, it is also necessary to involve the domain knowledge for managing those data.

Conventionally, spatial information is represented largely in a floor plan. A floor plan is well organized with a proper format and symbols and carries so many architectural meanings, from intuitive concepts to detail indications. So a content extracted from a floor plan is often regarded as the content of that space.

The spatial information represented in the floor plan includes spatial networks, which imply the unit nodes of spaces, and the linkages between them. The unit node of space represents the basic element of which a whole plan consists, i.e., a room could be a unit node of space in the whole house floor plan. The relationship represents spatial characteristics such as adjacency and connectivity between nodes. The spatial network is efficient for representing the matters of relationships. For example, it could be compared to a kind of frame containing the architectural domain knowledge.

Table 1. Metadata Classification

Classification		Applicable Metadata Item on Spatial Information
Content-independent Metadata		Address, Builder, Construction Method, Year Built, File Format, Scale of Drawing
Content-dependent Metadata	Non-semantic Metadata	Building, Plan, Zone, Space, Wall, Opening; Area, Ratio, Orientation, Number, Materials, Dimension; Household Characteristics, Duration of Stay, Passage Circulation, Use Condition,
	Semantic Metadata	Adjacency, Connectivity, Arrangement, Composition, Space Depth*, Degree of Symmetry, Degree of Condensation, Zoning, Period Trend, Satisfaction,

Classifying metadata to get suitable abstraction aids in exploiting metadata (Sheth, and Klas 1998). Because of the same goal - the content-based retrieval - the classification of multimedia metadata is a good mirror for that of spatial information. Much research has presented the classifications (Grosky 1994; Bohm and Rakow 1994; Boll et al. 1998; and Klippgen et al. 1998), but there is a tendency to divide metadata into two main classes, content-dependent and content-independent metadata. The content-dependent metadata roughly has also two branches: to have semantics or not to have. Semantics is the meaning involved by knowledge, human perception or cognition. Table 1 shows the metadata classification and application to space generally.

3.2 Classifying Spatial Information of an Apartment Unit Floor Plan

There has been much research representing diverse viewpoints, associated with an apartment unit floor plan (Kim and Park 1992; Jang and Choi 1994; Cho and Park 1998; Park and Ju 1999; Han and Choi 2001). It is possible to extract the factors to determine spatial plan from the research. They can be assigned to four categories: general substance, dweller's behavioral substance, physical substance, and the inferable substance. The general substances and behavioral substances can be extracted from statistic materials or surveys, the physical substances from a floor plan (drawing), and the inferable substances would be the investigated result of synthesizing the other factors. All factors in these categories could individually be metadata items. Related to these categories, we can forecast some kinds of queries for retrieving the data. Table 3 shows the classification of spatial information associated with an apartment unit floor plan and query examples. These are some examples of the many possibilities for the spatial information search.

4. Data Model for Spatial Information

4.1 Concept of Structured Floor Plan

There has been some research on developing CAD systems that can build structured floor plans (Yessios 1986a, 1986b; Kalay et al. 1995; Choi 1997). In such systems, a floor plan is structurally well defined by having a hierarchical structure of components. Since space and form are the two main aspects in describing a building, they should be represented together in a system. To be more effective, this can take an object-oriented approach, where each building component is an object from the view of an object-oriented paradigm. The object has its own data and methods of how to behave in certain situations.

4.2 A Data Model for an Apartment Unit Floor Plan

Based on the building data model, a new data model can be developed for an apartment unit floor plan. It consists of two parts: classes for physical hierarchy and attributes. The physical hierarchy of spatial information, consists of, building, plan, zone, space, wall, opening, hole, door, and window. The attribute class consists of two parts: the general attributes (label, type, orientation, materials, number, dimension, ratio, area, and volume), and relational attributes (connected_to, adjacent_to, and facing).

Fig.1. Data Model for Space

5. Metadata for Spatial Information

5.1 Strategies to Generate Metadata

In order to generate metadata of space a spatial network consisted of the unit nodes of spaces and their linkages must first be defined.

Fig.2. Example of a Spatial Network for an Apartment Unit Floor Plan

A spatial network is a frame to describe spatial information. Vitruvius, a real-time floor-plan-structuring system, can automatically make a floor plan structured, just by drawing a floor plan. No matter how automatically it does, for example, in Figure 2, an apartment floor plan can be structured with several unit nodes. Within an apartment unit floor plan, a room or an individual space can be the unit node and their relationship can be the linkage.

5.2 Organizing Metadata

Resource Description Framework (RDF), as a foundation for processing metadata, provides interoperability between applications that exchange machine-understandable information on the web (W3C, 2000). It was developed by the World Wide Web Consortium (W3C) to create a standard format for representing and transporting information. RDF is an application of XML that imposes needed structural constraints to provide an unambiguous method of expressing semantics (Miller 1998). Additionally, the XML aspect is important for several reasons.

Using the language defined in the schema, we can describe the content of space on a floor plan. Figure 11 illustrates how RDF represents an example floor plan. An apartment unit floor plan is a Plan, and it would be divided with several unit nodes of spaces, which are Spaces in the physical data structure. An example of the relationship between spaces is shown in Figure 3.

5.3 Exploiting Metadata: Query Operation

The users specify the request with various query interfaces. By querying attribute predicates, the system applies exact-match retrieval, same as the traditional way. On the contrary, as for structural, content-based, and fuzzy predicates, the system makes the results a set of objects, each of which has an associated degree of relevance. They also need to process and optimize query for internal representation.

In the area of architecture, there is one more thing to be concerned on the interfacial issue. As the interface to query, the graphical environment is more suitable, while textual environment is more general in the IR field. It would be efficient for users in any period of retrieval, i.e., querying, answering, and even indexing.

A set of sample queries is presented to illustrate how to use query with knowledge. If a user asked a query of 'Find all plans having the semi-separated L-DK', the system should be able to accept the knowledge-layer metadata and ontology. That is, the system should understand what a semi-separated L-DK is like. In this case, it is necessary to prepare an ontology for relationships of L-D-K (Living room, Dining room, and Kitchen). Table 3 shows the definition and distinction of L-D-K composition (Choi 1998).

Fig.3. RDF Schema for the Physical Hierarchical Classes of Spatial Information

Table 2. The Classification of Relationships between L-D-K (Living room, Dining room, Kitchen)

Unification	Refraction
A living room, a dining and a kitchen are in an unified area.	A living room separated and a kitchen are in an partially from a dining and kitchen by refracting wall.
Semi-Separation	Separation
A living room is semi-separated from a dining and kitchen with a passage or a entrance.	A living room is separated entirely from a dining and kitchen.

Table 3. The Process of Distinction of L-D-K Composition

Stage 1: Check if a plan has spaces labeled as DK and L.
SELECT pln_id, spc_id, FROM sc_Space, sc_Plan WHERE (sc_Space.spc_label=DiningKitchen) and (sc_Space.spc_label=Livingroom)
Stage 2: Check if DK and L are abutting.
SELECT pln_id, spc_id FROM sc_Space, sc_Plan WHERE isAbutting ((sc_Space.spc_label=DiningKitchen) and (sc_Space.spc_label=Livingroom))
Stage 3: Check if the type of wall between DK and L is solid.
SELECT pln_id, spc_id, wl_id FROM sc_Space, sc_Plan, sc_Wall WHERE isSolid ((sc_Space.spc_label=DiningKitchen) and (sc_Space.spc_label=Livingroom))
Stage 4: Compare the dimensions among the spaces and partition wall.
SELECT pln_id, spc_id, FROM sc_Space, sc_Plan, sc_Wall WHERE (sc_Wall.facing=DiningKitchen) and (sc_Wall.dimension.lenth>maxWall(DK))
Stage 5: Check if the wall has openings.
SELECT pln_id, spc_id FROM sc_Space, sc_Plan, sc_Wall, sc_Opening WHERE (sc_Wall.haveOpening != 0)
Stage 6: Check the types of opening.
SELECT pln_id, spc_id, FROM sc_Space, sc_Plan, sc_Opening WHERE sc_Opening.door.type=001

6. Implementation of SpaceScope

SpaceScope is a spatial information retrieval system as a prototype system developed in this research. It consists of three parts; modeling, indexing, and searching.

Tabel 4. The Modules of SpaceScope

Module	Developing Environment	System
Modeling	Vitruvius (C++)	Vitruvius
Indexing	MS SQL Server 2000	SpaceCore
Searching	MS SQL Server 2000	SpaceScope Browser and Browsing XML, HTML, SQL

Modeling Module - Vitruvius

Vitruvius is a name of a real-time floor-plan-structuring system. No matter how automatically it does, an apartment floor plan can be structured with several unit nodes. Within an apartment unit floor plan, a room or an individual space can be the unit node and their relationship can be the linkage.

Indexing Module - SpaceCore

Based on the physical hierarchy of the data model, tables in the SpaceCore database are constructed. Building, Plan, Zone, Space, Wall, and Opening become tables, and their subclass and properties become columns in the tables. Each table has a primary key as a unique identifier, and the relationships between the tables are combined with foreign keys.

Browsing Module - SpaceScope browser

The SpaceScope browser will be executed on the web. In this research, only the browsing and searching module can be implemented on the web based on the HTML and XML connected to database. The interface for query specification is one of the most important issues to be discussed. Spatial information in architecture is sometimes very inefficient to be represented only by text, even though text is very easily computable. For example, if a user can make a query just by drawing a floor plan, it will be a new challenge in the architecture field.

Fig. 4. The Algorithm for Recognizing the Type of LDK

Fig.5. The System Architecture of SpaceScope

Table 4. Modeling Module, Vitruvius : Automatic Floor Plan Structuring System

Input Mode	Structure Mode
Input Mode deals with schematic drawings. Drawing lines, polyline, rectangle, or editing them. After finishing drawing, it can be converted into a structured floor plan.	Structure Mode deals with a structured floor plan. The drawings are not just a collection of geometric elements any more. They consist of building components such as space, wall, surface.

Diagram Mode	Model Mode
Diagram Mode deals with diagrams. Because the data structure is stored automatically, the possibility to draw diagram is very high, if there is the any logic to organize.	Model Mode deals with a 3-dimensional model. Currently, a MetaStream browser is used for the visualization.

Fig.6. The Interface of SpaceScope Browser

7. Conclusion

The potentials of spatial information retrieval also would be a big issue. Associated with the metadata and spatial information retrieval, this research may make conclusions with suggestion for such potentials as follows:

Constructing the floor plan database and content-based search engine This research suggests a novel approach to construct the floor plan database, based on the content of space. According to structured metadata describing the raw data object (a floor plan) semantically, the database can manage the spatial data efficiently in many architectural applications; also machinate higher layered information semantics or knowledge.

Analyzing the huge amount of spatial factors One of the most remarkable potentials shown in this research is the capacity to manage data. It covers the matter of not only quantity but also quality.

Estimating and evaluating the spatial information As for spatial information, this research suggests the concept of the spatial network, which materializes the intangible space as a data object. Based on this frame, spatial information can be estimated and evaluated in a computational way.

These conclusions may extend some expectations for related subjects as follows: 1) Industrial effect related to spatial matters, e.g. real-estimate, facility management, 2) Supporting potentials of case-based design

For the future works some have been also identified: it is necessary to construct XML-based databases and a knowledge based query language, such as RQL/XQL, working on such databases.

The more specific domain knowledge is involved, the more practical system would be. The domain knowledge may range diverse section even in architecture, such as

design, building performance, facility management, energy management, post occupied evaluation, historical research, and so on.

The interfacial issue (regarding the graphical visualization) should be investigated in depth for making query, or browsing answer, adequate to needs for the architecture field.

Acknowledgments

This work was supported in part by the Korean Ministry of Information and Communication under Information Technology Research Center (ITRC) Program.

References

- 1) Baeza-Yates, R. and Ribeiro-Neto, B. (1999) Modern Information Retrieval. New York: ACM Press, Addison-Wesley.
- 2) Bray, T. (2001) XML.com: What is RDF?
- 3) <http://www.xml.com/pub/a/2001/01/24/rdf.html> page=1
- 4) Bohm, K. and Rakow, T.C. (1994) Metadata for Multimedia Documents. SIGMOD Record 23(4): 21-26.
- 5) Boll, S., Klas, W. and Sheth, A. (1998) Overview on Using Metadata to Manage Multimedia Data. In Multimedia Data Management, ed. Sheth, A., and Klas, W. New York: McGraw-Hill
- 6) Chien, S-F. and Shih, S-G. (2001) Design through Information Filtering. CAAD Futures 2001 Proceedings: 103-110. Netherlands: Kluwer Academic Publishers
- 7) Cho, S-H. and Park, S-B. (1998) A Study on the Users' Need for Developing Diverse Dwelling Unit Plans of Apartment. Journal of the Architectural Institute of Korea 14(11).
- 8) Choi, J-W. (1997) A Development of an Intelligent CAD Engine to Support Architectural Design Collaboration, Korea CAD/CAM Journal: 53-59.
- 9) Gill, T. (2000) Introduction to Metadata: Metadata and the World Wide Web <http://www.getty.edu/gri/standard/intrometadata>
- 10) Hiller, B. and Hanson, J. (1984) The Social Logic of Space. London: Cambridge University Press
- 11) Hjelm, J. (2001) Creating the Semantic Web with RDF: professional developer's guide. New York: John Wiley & Sons, Inc.
- 12) Jang, S-J. and Choi, D-W. (1996) Problems of Convex Space Drawing in Space Syntax; Cases of 4 Model Houses with Concave Space, Journal of the Architectural Institute of Korea 12(4), 43-51.
- 13) Kalay, Y.E. Khemlani, L. & Choi, J. (1998) An integrated model to support distributed collaborative design of buildings, Automation in Construction (7): 177-188.
- 14) Kim, C-H. and Park, T-Y. (1995) A Study on the Process of Layout Planning and Characteristics of Computer-aided Space Planning. Journal of the Architectural Institute of Korea 11(7): 41-53.
- 15) Kim, S-A. and Park, Y-H. (1992) Inflexible Patterns of Apartment Unit Plan: The Comparative Analysis of the Public and Private Sector Apartment. Journal of the Architectural Institute of Korea 8(7)
- 16) Miller, E. (1998) D-lib Magazine: An Introduction to the Resource Description Framework
- 17) <http://www.dlib.org/dlib/may98/miller/05miller.html>
- 18) Sheth, A. and Klas, W. (1998) Multimedia Data Management. New York: McGraw-Hill
- 19) Yessios, C.I. (1986) The Computability of Void Architectural Modeling, The Computability of Design, Proceedings of Symposium on Computer-Aided Design at SUNY Buffalo, New York
- 20) Zhou, Q., Krawczyk, R.J. and Schipporeit, G. (2002) Re-searching the Research Problem in CAAD. Proceedings of CAADRIA 2002 : 73-80. Malaysia: Prentice Hall