SpaceScope: Spatial Content-Based Retrieval of Architectural Floor Plans

Jie-Eun HWANG and Jin-Won CHOI
Department of Housing & Interior Design, Yonsei University, Korea

Keywords: case-based reasoning, content-based retrieval, metadata, RDF, semantic

Abstract: This research investigates spatial information retrieval in architecture, based on an efficient metadata that is crucial for content-based retrieval. SpaceCore is the metadata schema for spatial information on a floor plan implemented with RDF (Resource Description Framework). To generate metadata, we analyse the factors of spatial information, and then construct a data model containing the content and structure of spatial information. To exploit the metadata, we suggest various query operations with possible predicates and interfaces. SpaceScope, the spatial information retrieval system, consists of three modules: an input data module, a query module, and a browse module. As to implementation, we have already developed the real time plan-structuring system, called "Vitruvius", as an inputting and modelling data module (using C++). Based on the database, the system performs the similarity analysis and then evaluates the alternative plans with the retrieval goals. At the end of the paper we discuss the potential of SpaceScope in the schematic design stage of the architectural design process, possibly improving and triggering the designer's conceptual approach against a logical background and facilitate rapid decision-making after full consideration.

1 INTRODUCTION

In the architectural design domain, information retrieval (IR) has been oriented empirically since case-based design (CBD) began to be discussed as an important area in computational design (Flemming 1994; Rosenman 2000). Despite its powerful potential in the design process, CBD has faced such problems as the representation of design cases, indexing, matching, retrieval and selection, and adaptation in a result (Rosenman 2000), impeding pragmatic application for design.

As for the data concerned in the architectural field, including aspects outside of design matter, the research up to this point has tended to incline toward questions of how to convert conventional data into digital form efficiently or how to standardize such data. Most of all, although space is fundamental in architecture, the majority of research only mentions the enclosures of spaces, such as building components - wall, floor, roof, equipment, etc. - represented in drawings, specifications (Zhou et al. 2002), or the external factors about the building - the building name, the building type, the related references- (Akin 2002). Further, the research concerned directly with spatial matter, for example, Space Syntax (Hillier and Hanson 1984), did not treat a large amount of general cases. There has been limited research on spatial information, especially for the huge capacities that would accommodate a case base.

Recently, the scope of IR has expanded far beyond its original purpose, such as indexing and searching for useful documents. The huge amount and various types (e.g. multimedia) of information has accelerated to 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). Consequently, IR includes classification and categorization, systems architecture, user interfaces, data visualisation, filtering, languages, etc.

The approach of content-based retrieval can be a good solution to overcome the limitation of the management of spatial information in the design process. Content-based retrieval is currently a core issue for IR, especially for multimedia IR. There has been satisfactory research that is adaptable to spatial information in terms of both technical and conceptual issues, given the similarity of characteristics of multimedia and architectural data.

This research investigates spatial information retrieval in architecture based on efficient metadata that is crucial for content-based retrieval. Concentrated on the design stage, we present how to make and manage the metadata - the retrieval strategies.

2 THE NEED FOR CONTENT-BASED RETRIEVAL

The case approach regards the design process as learning from precedents and solving practical problems through those referents. There may be no wonder about this aspect of CBD, since no designer in the world could do without precedents. However, the essence of the case method is the presentation of problems through past cases and the context documented in written form. A case base, therefore, is the collection of instances or cases usually codified in a manual or computational database (Akin 2002). Although it corresponds to the current technical backdrop, the reason that the case-base approach has not overcome the problem of pragmatic application is primarily the insufficiency of the IR method.

3 METADATA FOR SPATIAL INFORMATION

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 organisation of data, the various data domains, and the relationship between them (Baeza-Yates and Ribeiro-Neto 1999). Demands for metadata have grown rapidly, especially in an area such as multimedia IR, because of the need for a different querying paradigm, the adequate and efficient retrieval processing techniques, and the semantics of multimedia data (Boll et al. 1998). Similar to recent multimedia information retrieval approaches, a spatial IR system should be able to afford users the ability to specify not only data attributes but also an object's content.

3.1 Metadata and Content-based Retrieval

Classifying metadata to identify the semantics assists in exploiting metadata (Sheth and Klas 1998). Much research has presented the classifications (Grosky 1994; Bohm and Rakow 1994; Boll et al. 1998 and Klippgen et al. 1998). In prior research, there has been a tendency to divide metadata into two main classes: content-dependent and content-independent metadata. The content-dependent metadata has roughly two branches: with semantics or not. Table 1 shows the metadata classification and its general application to space.

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	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,

* Space depth is the value indicating how far the room is to enter from the other rooms in a whole plan

3.2 A Multi-layered Hierarchy Metadata Model for Representing Spatial Information

Multi-layered metadata model (Chu, W.W. et al. 1998) is an adaptable approach for spatial information retrieval. It was suggested for a medical picture archival and communication system to provide content-based retrieval. Since it is important to retrieve the diagnostic images by content within historical cases in therapy treatment planning, metadata should be sophisticated enough to capture the content efficiently. It includes three layers: 1) Representation Layer for providing the characteristics of the raw images and pixel-level description of the image content, 2) Semantic Layer for storing domain-dependent object features and spatial relationships, and 3) Knowledge Layer for representing the knowledge of spatial, dimensional, temporal and evolutionary characteristics of image content.

The similarities between medical image retrieval and spatial information retrieval can be deployed on: 1) concerning pictorial data, 2) the importance of analysing spatial relationship, and 3) the need for involving particular domain knowledge. Therefore the concept of a multi-layered data model is basically appropriate to identify the spatial data. There is, however, a striking diversion between them in the structure of data. Architectural pictorial data does denote hierarchical meanings above pixel-level description. For instance, a line representing a wall symbolizes not only the existence of the wall and its attributes, but also the context, such as the spatial relationship divided by the wall, the roll for structural mechanics in an entire building, the opening type it affords, and so on. For those reasons, we define the metadata model through a hierarchical framework and divide it into three layers and four groups of classes in the retrieval phases: representation layer, semantic layer, and knowledge layer.

In this research, we are primarily concerned with the Semantic Layer - the relationships between spaces. It will be sufficient, however, to illustrate how to work the metadata in the retrieval system with semantics and domain knowledge.

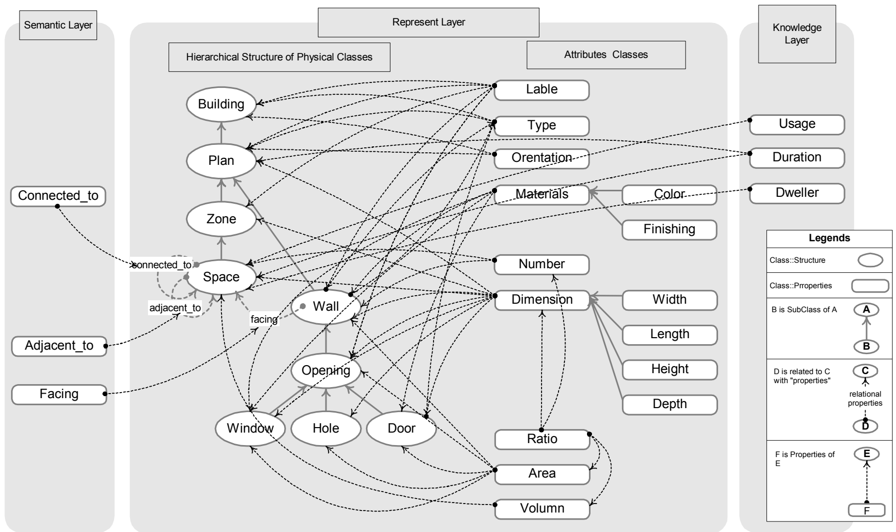
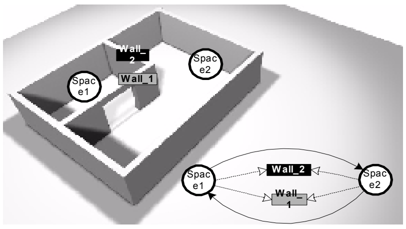


Figure 1 A Multi-layered Metadata Model

3.3 SpaceCore: RDF Schema for Spatial Information

Resource Description Framework (RDF) was developed by the World Wide Web Consortium 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.

SpaceCore is an RDF schema describing the structure and framework of spatial information to afford the diverse types of retrieval. SpaceCore consists of four parts, as reflex of a metadata model: 1) physical hierarchical classes, 2) general attributes classes, 3) relational attributes classes, 4) properties of each physical hierarchy. Figure 2 shows the schema for the physical classes of spatial information.




```
<?xml version="1.0"?>

<rdf:RDF xmln:rdf="http://www.w3c.org/1999/02/22-rdf-syntax-na#"
         xmln:rdfs="http://www.w3c.org/2000/01/rdf=schema#"
         xmln:sc="http://archmedia.yonsei.ac.kr/SpaceScope/SpaceCore#">

<rdf:Plan ID="Plan_1">
   <rdf:Space ID="Space_1">
      <sc:adjacent_to> Space_2 </sc:adjacent_to>
      <sc:have_wall> Wall_1 </sc:have_wall>
      <sc:have_wall> Wall_2 </sc:have_wall>
   </ref:Space>

   <rdf: Space ID="Space_2">
      <sc:adjacent_to>Space_1</sc:adjacent_to>
      <sc:have_wall> Wall_1 </sc:have_wall>
      <sc:have_wall> Wall_2 </sc:have_wall>
   </ref:Space>

   <rdf:Wall ID="Wall_1">
       <sc:wl_type>solid</sc:wl_type>
       <sc:have_opening>Opening_1</sc:have_opening>
       <sc:face_on>Space_1</sc:faced_on>
       <sc:face_on>Space_2</sc:faced_on>
    </rdf:Wall>

   <rdf:Wall ID="Wall_2">
       <sc:wl_type>solid</sc:wl_type>
       <sc:face_on>Space_1</sc:faced_on>
       <sc:face_on>Space_2</sc:faced_on>
    </rdf:Wall>

   <rdf:Opening ID="Opening_1">
       <sc:op_type>door</sc:op_type>
    </rdf:Opening>

</rdf:Plan>

</rdf:RDF>
```

Figure 2 The schema for the physical classes of spatial information, the example representation.

4 RETRIEVAL STRATEGIES

In this research, we exemplify the traditional folk house floor plan database to demonstrate the retrieval strategies.

4.1 Overview of Spatial Retrieval System

A spatial IR system should be able to deal with a complexly structured object supported by the underlying data model, the query language, and the mechanisms of access and storage (Baeza-Yates and Ribeiro-Neto 1999). The architecture of SpaceScope is illustrated in Figure 3. There are roughly three process modules:

- Input data module: Input data modelling metadata with the schema
- Query module: Graphical query
- Browse and evaluate module: Intelligent browse.

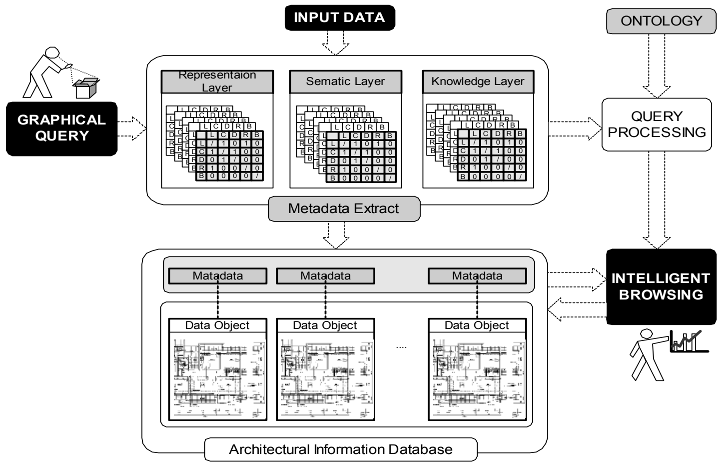


Figure 3 The Architecture of SpaceScope

4.2 Input Data

4.2.1 Vitruvius

There has been some research on developing CAD systems that can build structured floor plans (Kalay et al. 1995; Choi 1997). In such systems, a floor plan is structurally well defined by having a hierarchical structure of components. Vitruvius is the name of a real time floor-plan-structuring system. Simply drawing a floor plan in Vitruvius, a user can get metadata from the original data object, constructed by the SpaceCore RDF specification.

4.2.2 An Ontology of Spatial Relationships

Ontology is a specification of conceptualisation. Ontology defines the terms used to describe and represent an area of knowledge. Ontologies include computer-usable definitions of basic concepts in a domain and the relationships between them.

The data in the semantic layer and knowledge layer is organised on ontologies providing the conceptual terms and query predicates. For example, between two spaces, there are three types of relationships - separate, adjacent, and connected. A ‘connected’ type is divided into three types according to the types of wall and opening - connected with hole (virtual wall), door, and window (Figure 4). In this way, an ontology of spatial relationships categorizes the metadata in a semantic layer as well as the predicates for query. By extending specific domain knowledge - for example, in terms of typology for dwelling space - we can develop an ontology for the relationship between a kitchen and a living room. These may be represented through such conceptual word as separation, semi-separation, unification, and refraction.

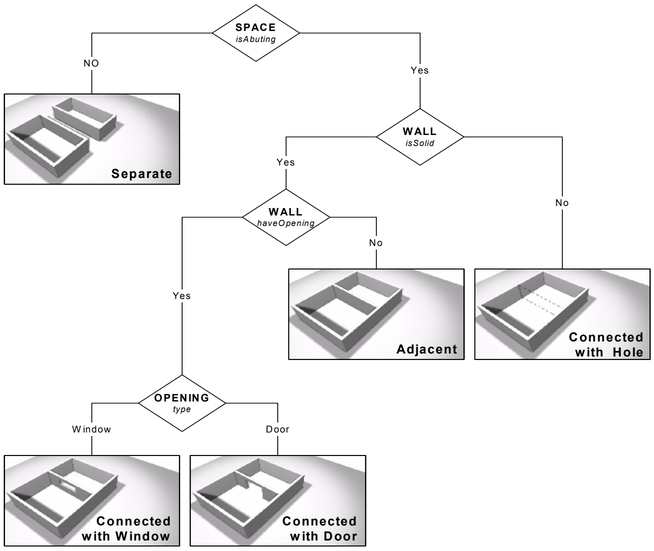


Figure 4 An Ontology of Spatial Relationship

4.3 Graphical Query

Users specify a request with various query interfaces. Below is an example of the predicates applied to spatial information (Baeza-Yates and Ribeiro-Neto 1999):

By querying attribute predicates, the system applies exact-match retrieval, just as in the traditional IR way. In contrast, the system, concerning structural, content-based, and fuzzy predicates, makes of the results a set of objects, each of which has an associated degree of relevance. They also need to process and optimize queries for internal representation by ontologies.

The most novel as well as difficult aspect of spatial information retrieval is the method of query. Due to the fact that the majority of a process in a design project life cycle proceeds in a graphical way - sketch, drawing, rendered image, and so on - it is sometimes very abstruse to convert it into a textual version.

In fact, the internal process of query would work with the textual format, but it is not always easy to convert graphical thought into the textual version, even if the WHERE clause can be replaced by the simple predicates based on an ontology.

Figure 5 is the intuitive interface for graphical query. A user can draw a plan in Vitruvius and save it as a schematic plan complying with the metadata schema, and then open it in the graphical query window. The option box containing the query predicate would be diverse in developing the method to ménage metadata in the

Knowledge Layer with particular domain knowledge.

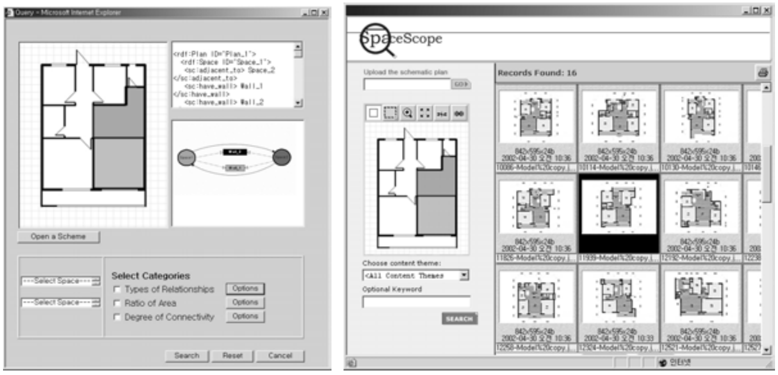


Figure 5 Graphical Query Windows

4.4 Intelligent Browsing of Candidates

After a query processing, a user should browse candidates and make the decision to adapt. All candidates embrace the user's goal to retrieve (or design), so the system should log the retrieval procedure and visualise the goal. As soon as a user chooses the selection, she/he can contact the original data object. As for the Figure 5 query, the list of candidates provides those that have a similar relationship between two spaces to the query plan in terms of the degree of connectivity and the area ratio. In addition, for adaptation to the design, the logging data will be attached in the database as a procedural history of design.

When SpaceScope contacts other databases which convey the outer factor data related to the floor plan, it could perform more intelligent browsing. As an instance, the location-based traditional folk house database can be exploited on the analysis the relationship between the spatial composition and the locality. Connecting the geographic data with the spatial information of floor plans, we can examine the further fundamental and intangible determinant factors of the design such as landforms, climates, vegetations, populations, and social environment. In terms of a traditional Korean folk house, there has been a lot of research mainly in the viewpoint of typology, paying attention to elements to determine the type of a plan. Along with SpaceCore schema, we defined the elements of a folk house floor plan, and then can add the location data within the database. To investigate the relationship between a folk house floor plan and locality, we can apply the GIS technology on it. The system will convey the denoted meaning of a floor plan with innovative computational method. The statistics has been considered as the best way to examine the relevancy among some factors throughout previous research. However, we can develop a novel method to carry out a thorough investigation of design determinants. Also based on the prior research, we can employ SpaceScope for practical adaptation in design development.

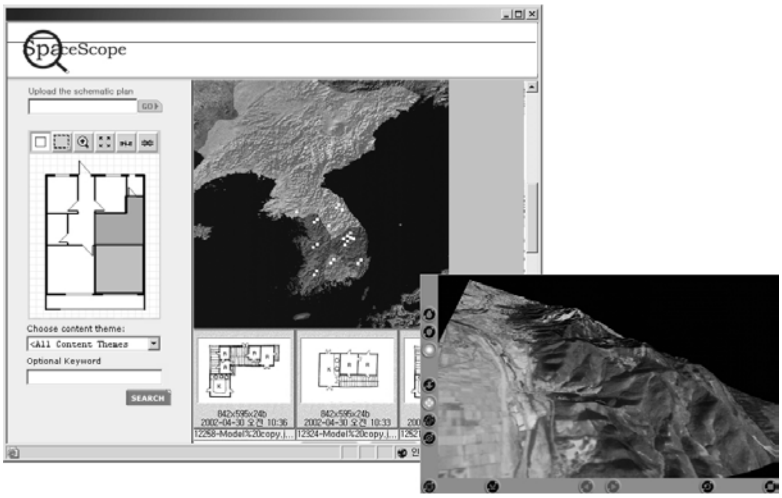


Figure 6 Browsing Candidate Floor Plans with Geographic Factors

5 CONCLUSION

SpaceScope is a content-based retrieval system that deals with architectural floor plans. In terms of its contribution to the design domain, metadata and spatial information retrieval, we suggest that it contains the following potential:

1) Constructing the floor plan database with semantically rich metadata.

2) Managing and analysing a 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. Rapid processing, broad analysis, and explicit synthesis can be executed in an unprecedented way.

3) Proposing an innovative query interface appropriate to design process.

Some future works have also been identified:

- In terms of case-based design, evaluating the candidates and adapting to the design intention should be investigated.
- The domain knowledge may expand to diverse sectors to include even architectural issues such as design, building performance, facility management, energy management, post occupied evaluation, historical research, and so on.
- To implement more empirically, it is necessary to construct XML-based databases and develop a knowledge-based query language such as RQL/XQL.

ACKNOWLEGEMENTS

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

REFERENCES

Akin, O. 2002. Cased-based Instruction Strategies in Architecture, *Design Studies*, 23(4). London: Elsevier Science Ltd.

Baeza-Yates, R. and B. Ribeiro-Neto. 1999. *Modern Information Retrieval*. New York: ACM Press, Addison-Wesley.

Boll, S., W. Klas and A. Sheth. 1998. Overview on Using Metadata to Manage Multimedia Data. *Multimedia Data Management*, eds. A. Sheth and W. Klas, 1-24. New York: McGraw-Hill.

Choi, J.W. 1997. A Development of an Intelligent CAD Engine to Support Architectural Design Collaboration. *Korea CAD/CAM Journal*, 53-59.

Chu, W.W., C.C. Hsu, I.T. Ieong and R.K. Taira. 1998. Content-Based Image Retrieval Using Metadata and Relaxation Technique. *Multimedia Data Management*, eds. A. Sheth and W. Klas, 149-190. New York: McGraw-Hill.

Flemming, U. 1994. Case-based Design in the SEED System. *Knowledge-Based Computer-Aided Architectural Design*, eds. G. Carrara & Y. Kalay 69-91. Amsterdam: Elsevier Science Ltd.

Gill, T. 2000. Introduction to Metadata: Metadata and the World Wide Web. http://www.getty.edu/gri/standard/intrometadata

Hiller, B. and J. Hanson. 1984. *The Social Logic of Space*. London: Cambridge University Press.

Hjelm, J. 2001. *Creating the Semantic Web with RDF: Professional Developer's Guide*. New York: John Willey & Sons, Inc.

Rosenman, M. 2000. Case-based Evolutionary Design. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, 14(1): 17-19. USA: Cambridge University Press.

Rivard, H. and S.J. Fenves. 2000. SEED-Config: A Case-based Reasoningsystem for conceptual building design. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, 14(5): 415-430. USA: Cambridge University Press.

Zhou, Q., R.J. Krawczyk and G. Schipporeit 2002. Re-searching the Research Problem in CAAD. *Proceedings of CAADRIA 2002*, eds. A. Eshaq, W.K. Chee, K.T.K. Neo, M. Neo, S.N.S. Ahamad, 73-80. Selangor, Malaysia: Prentice Hall.