Indexing and Retrieval of Visual Design Representations

Alexander Koutamanis[1], Gilles Halin[2], Thomas Kvan[3]
[1]Delft University of Technology, The Netherlands, [2]Nancy University, France, [3]University of Melbourne, Australia
[1] http://www.re-h.nl, [2] http://www.crai.archi.fr/integration/index.php, [3] http://www.abp.unimelb.edu.au/dean/
[1] a.koutamanis@tudelft.nl, [2] gilles.halin@crai.archi.fr

Indexing and retrieval of architectural visual databases refer to multiple levels of abstraction and various points of view which may co-exist in a single image. This complexity is increased by the necessity to structure architectural images into well-defined, meaningful representations. We propose that the correlation of domain analysis and general methods and techniques provides the background to the solution of most problems and a deeper understanding of the structure of indexing and retrieval in architecture.

Keywords*: Representation; database; indexing; retrieval.*

Introduction

In a heavily networked world of increasingly digitized information, the utility of architectural images grows in importance. This has to do with the preservation and dissemination of historically important documents (a subject beyond the scope of the present paper) but also with certain demands of architectural design. We use digital architectural images for a number of purposes, from the synchronous and asynchronous exchange of design information to referring to precedent design solutions and to describing physical or social contexts in a non-committal way. The success of architectural images relies primarily on our cognitive capabilities which allow us to recognize and identify relevant entities and relations, and reason on the basis of information that goes beyond what is explicitly conveyed by the image. We seldom consider such basic processes with the awe they deserve but empowering the computer with similar capabilities is not easy. Current computers do not possess our tremendous capacities of perception and memory (Kosslyn, 1994) that make recognition and identification seem so effortless.

Uses of architectural images that form the subject of the present paper concern the related needs of two different parties. On the one hand we have the user's need to find sufficient relevant and appropriate information in one or more images and on the other the producer's (or distributor's) need to offer such information in a way that not only makes it useful but also unambiguous (from the producer's or distributor's point of view). Coupling the two is normally based on methods for organizing information in manners that enhance usability and reduce the number and magnitude of possible errors. Using such methods the producer's or distributor's basic problem is resolved by means of indexing, i.e. associating an image or its parts to comprehensive, transparent structures of relevant concepts. Such associations translate any image into a collection of terms that describe salient features and allow for classification with respect to a number of aspects. These include the type of the image, its subject, morphological and typological information, spatial and structural information, as well as performance characteristics. These terms also support resolution of the user's basic needs by allowing transparent matching of any issue that may be of interest to available information, from general, pragmatic requests to very specific requirements relating to particular aspects and actions'. By translating such requests for information into queries comprising the same terms used in indexing we can support effective, efficient and reliable retrieval from a variety of viewpoints.

Methods and techniques for indexing and retrieval pre-date computerization and the Internet. Even though they were given new impetus and rigor by computer science, many derive from various disciplines among which library science arguably predominates. Most methods were initially oriented towards textual documents and information (e.g. textual databases, books and journal articles). This bias is still evident and relates to one of the central problems in architectural representation, the difference between propositional and depictive representations.

Recent developments

Even before the emergence of the Internet, the availability of optical digitizers, multimedia systems and affordable computer storage led to many attempts at transferring existing images from analogue carriers to the computers. The results have been variable, both technically (i.e. with respect to image quality) and organizationally (with respect to database structure, indexing and retrieval). Research into the structure of visual architectural databases has also been motivated by different approaches to automated design which relied on the use of precedents, types and prototypes (Aygen and Flemming, 1998, Brown and Steadman, 1986). Such research resulted in the identification of indexing and retrieval strategies that relate to design processes, including the capacity of formalisms used in generative systems to signify relevant aspects and parts in an image (Koutamanis, 2000). A related line of investigation concerns analyses of architectural images as sources of information that can be recognized and interpreted (Bignon et al., 2000, Gross, 1995, Kacher et al., 2002).

Probably the highest point of research into the indexing and retrieval of architectural images has been the Getty Art & Architecture Thesaurus (AAT), an extensive collection of domain indexing terms into a comprehensive controlled vocabulary for describing concepts in architecture and the visual arts (http://www.getty.edu/research/conducting_research/vocabularies/aat/: May 2007). The importance of the AAT lies not only in the integration and codification of existing terminologies but also in that it opened up possibilities for a consistent treatment of architectural information in databases of images (Porter and Thornes, 1994).

Despite such advances, most visual architectural databases remain quite modest in terms of indexing and often fail to exploit the support advanced indexing structures offer to query refinement, relevance feedback and other critical aspects of retrieval. Similarly, most image collections on the Internet and search engines tend to be rather basic with respect to indexing and quite rudimentary in terms of retrieval. Such limitations are significant in the case of digitized images but even more restricting when applied to the structured representations we produce with the computer. Computer-based representations are arguably developed with explicitness of parts and aspects in mind, so as to support more effectively and reliably design synthesis, analysis and communication. These are complemented by information standards developed with integration, continuity and interoperability in mind (Koutamanis et al., 2007). If we fail to match such structures with the capabilities of indexing methods and the needs and processes of retrieval then we fail to (a) extend the utility of representation and standardization schemes to related application areas like information management, communication and collaboration, and (b) compare and test existing methods and techniques developed from different viewpoints for the execution of related tasks.

Capabilities and requirements

A starting point to such integration is an overview of requirements from architectural processes as well as representations from the viewpoint of information utility. This overview reveals goals, priorities and obstacles in architecture and building, and facilitates links to approaches to the indexing and retrieval of visual representations, with particular emphasis on forms of search and query formulation.

Architectural representations

Computerization has brought on a radical change in the way we approach architectural representation. The immediate effect has been a transfer of analogue drawing practices to the computer. Change was primarily restricted to the implementation level, replacing the single stroke and the trace it leaves on paper with a similarly shaped graphic element. Despite the efforts of CAAD to add structure to architectural representation this basic change remains the main departure in design computing because it underlies the main practical advantages of digital representations like compactness, adaptability and transformability. The highly structured representations proposed in CAAD have yet to have an extensive effect, possibly because of their strong associations with particular generative systems (e.g. rectangular arrangements and shape grammars).

More appealing to practice have been attempts to organize strokes into descriptions of building elements and components (libraries). These attempts owe more to ideas about building industrialization than design computing (Koutamanis et al., 2007). Consequently they tend to focus on the last stages of designing and the transition to construction. Recent interoperability standards pay more attention to procedural aspects and changes of aspect, even though the transformation of spatial and building elements with changes of abstraction level (resolution) remains underdeveloped. Similar problems are encountered at the implementation level, as the composition of symbols out of strokes remains arbitrary (solely based on appearance).

Architectural perception

The focus of the imagery debate in cognitive science has been the difference between depictive and propositional representations (Kosslyn, 1994). Depictive representations are pictorial, conveying meaning through their resemblance to an object and consisting of parts that can be defined arbitrarily and flexibly. Propositional representations consist of symbols indicating relations (predicates) and objects. For example, ON (BEAM, POST1, POST2) is a propositional description of the basic trabeated module. CAAD had its own version of the same debate, even though it was quite one-sided. For a long time (especially in the 1970s and 1980s) CAAD underestimated the importance of drawings as depictive representations, continually stressing that 'D' stood for design and not for drawing (Kvan, 2004).

The transfer of architectural drawing to the computer was therefore either ignored (and left to commercial developments) or pushed towards a propositional direction that matched mainstream approaches to automated design. The propositional bias in CAAD may also relate to the character of the new implementation environment: in vector environments each graphic entity is discrete and identifiable. A similar propositional structure may have seemed obvious also for architectural entities. Still, there have been few attempts to bridge the gap between the propositional CAAD tendencies and depictive practices in an intelligent and relevant manner (Yessios, 1987).

From this perspective it is not surprising that interoperability standards and higher-level communication are dominated by propositional tendencies, even thought these may not fit the structure if either drawing or reality. It is doubtful that we draw as consistently as assumed in most approaches to standardization, while the relation between drawing actions and predefined symbols can be more complicated than selecting and manipulating a discrete entity. Similarly the decomposition of a real building into elements and components may result into variable configurations, alternative interpretations and overlapping entities (especially when we take into account the materials or techniques used and the different actors involved in construction). It is noteworthy that perceptual constraints play an important though understated role in this decomposition (Koutamanis, 2006). Equally understandable is that computer users have reacted with amorphous, unstructured or even unfeasible forms that are difficult to categorize.

The symbolic abstraction of propositional representations is highly attractive when we want to encapsulate complex information into economical concepts, recognize types or ignore trivial information. It allows for a higher level of reasoning that is unencumbered by concerns that can be deferred to a later moment or ignored altogether. We can think of A as a letter, as part of a sequence (first), as a sound, as an initial of a person's name etc. without having to consider the strokes it is composed of, the font it is printed in, its size, colour or location on a page. Similarly, a free-standing column can be treated as an abstract entity: a structural support, an obstacle to pedestrian movement, a source of shadow. Consequently it can be abstracted into a volume, a small collection of surfaces or even a vertical axis. We do not have to classify it as an Ionian, Corinthian or a Modernist stylistic element for such tasks.

Requirements

Information retrieval is one of the fundamental needs of architectural design. Unfortunately it might be too fundamental to become a practical or research priority. On the one hand, we are used to predefined structures (from information standardization schemes to design process specifications) in order to organize design communication and presentation. On the other, we have become accustomed to the rather sloppy performance of popular search engines. By any measure the precision (ratio of number of relevant documents retrieved to the total number of documents retrieved) and recall (ratio of number of relevant documents retrieved to the total number of available relevant documents) are not impressive.

Further complications are caused by the fuzziness of the context and the goals of architectural searches. Admittedly, the possible common departures and intentions of a general, encyclopaedic search for images of a precedent and a search for information on a building component (product) can be confusing as to the character of the search but there are two general criteria that suffice for categorization:

1. *Abstraction*: searches relating to general aspects of a design (as e.g. most precedent searches) refer to a high abstraction level characterized by principles of form or construction, diagrammatic descriptions of spatial arrangement or functional patterns, and general performance characteristics. Queries at this high level are generally hard to formulate with any accuracy and generally rely on search intermediaries and relevance feedback. At the other end of the spectrum we encounter searches for very specific parts of a design, similarly motivated by formal, structural or performance issues. The main difference lies in that the low levels can be described by explicit constraints from the current state and form of a design. These constraints facilitate the formulation of clear, effective queries. The spectrum between these two extremes is populated by combinations of abstract characteristics and patterns with concrete components and constraints. The abstraction level normally rises with the size of the part.
2. *Usability*: a search is the means for identifying appropriate information sources for solving a problem. Consequently, information retrieved by a search should become part of a design (including its reference frames). The most noncommittal use of retrieved information is simple as-

sociation with a part or aspect of a design. Precedents usually function in this manner, providing implicit justification for design decisions and choices. The same applies to contextual information, which usually serves as background to e.g. programmatic constraints and requirements. The most concrete form of use of retrieved information is integration in a design, e.g. inclusion of a product symbol from a provider's online library in a CAD file. In between association and direct integration we encounter various forms of transformation which match retrieved information to the format and constraints of the design documentation.

Indexing

The choice of a method that can be used to index and manage a collection of images depends primarily on the size of the collection. In small collections each image can be analysed manually but larger collections require automated methods. The techniques used in the domain of content-based image retrieval (CBIR) are based on the automatic extraction of the lowest level features such as colour, texture, shape and spatial location (Del Bimbo, 1999). Many CBIR systems, which has been developed since the early 1990s, use these techniques (Veltkamp and Tanase, 2002) and give relevant results when the collections used contain very specific images such as fingerprint and face representations or very broad images such as photographs of landscapes and paintings. However, they become irrelevant when the collections are heterogeneous (Kacher, Bignon et al. 2002).

The heterogeneous character of a collection makes the object recognition techniques implemented in an image corpus ineffective. Some systems are oriented towards shape-based retrieval on 2D or 3D images, but they are in an experimental form and applied on very precise types of collections (Vleugels and Veltkamp, 2002). To improve the relevance of the CBIR methods on heterogeneous collections of images, the contribution of the image meaning is necessary. Emergent statistical methods, as in the ALIP system, use machine learning techniques and statistical models to train the system to learn specific concepts in order to characterize it in low level features (Li and Wang, 2003). Other methods, applied on the web, use the context of the image to identify concepts. The context can be only the image tag in the HTML page, as in the Google engine or the text close to the image (Nakapan et al., 2004). For a higher conceptual level, a manual or semi-automatic method has to be used.

Semantic indexing approaches rely on an image analysis to identify the concepts present in the image. The semantic description of the image content can be in the form of an ordered or weighted list of keywords or could be a more structured representation as a graph known as "conceptual" (Polovina and Heaton, 1992). The semantic web and its standards (OWL, RDF etc.), the multimedia standard format, MPEG 7, and the underlying XML language give new perspectives to the semantic methods. The knowledge of a domain can be described inside an ontology, exchanged or shared inside a community, and used to index the content of a collection of images (Mezaris et al., 2003). The MPEG 7 format makes it possible to encode the image and their meta-data in order to be used by several systems inside a community or in a personal way. Emerging new systems use these standards to propose "semantic annotation", image sharing, and exploit the XML structure to evaluate the semantic relevance of an image (Lux et al., 2003).

Retrieval

Query mechanisms play a vital role in bridging the semantic gap between users and retrieval systems. CBIR systems usually employ Query By Example (QBE) interfaces. These interfaces let the user provide an example image that can be used for comparison by the system. This can be done by means of selecting an image that is already present in the database or by uploading a new image. The system can then look for images with similar features that will be presented to the user. This kind of interface is also used in semantic retrieval, as it allows the user to express his needs with means other than words. In some QBE interfaces, the user can give relevance feedback to the system by interacting with the retrieved selection visualized in an image mosaic form. He can point out which images are relevant to him and which are not. Next, he re-submits the query, refining his search and (usually) narrowing the diversity of the retrieved image collection. In semantic approaches, the relevance feedback method can be considered as a machine-learning process, where the system tries to represent (learn) the user needs from positive examples (relevant images) and negative ones (irrelevant images) (Halin et al., 1990). The ImageGrouper system proposes an alternative interface, which uses positive and negative examples (Nakazato et al., 2003). The user navigates inside the image collection with a QBE interface and builds progressively groups of images which can be considered as positive and negative. These groups can be annotated, kept in albums and re-used for new research.

Other CBIR systems let the user query for an image by sketching a drawing. Although this seems to be an interesting interface, it is hard to use for people who can not represent a real object with a pen on paper, let alone by using a mouse. It is also understandable that query-by-sketching is hard to use when the images in the database contain details that are difficult to reproduce by hand. Rather than drawing, the user can attract the system's attention to particular image aspects such as the spatial composition of desired content in terms of particular regions.

In order to obtain a more precise retrieval based on high-level features, some approaches propose a more expressive query formulation using words or

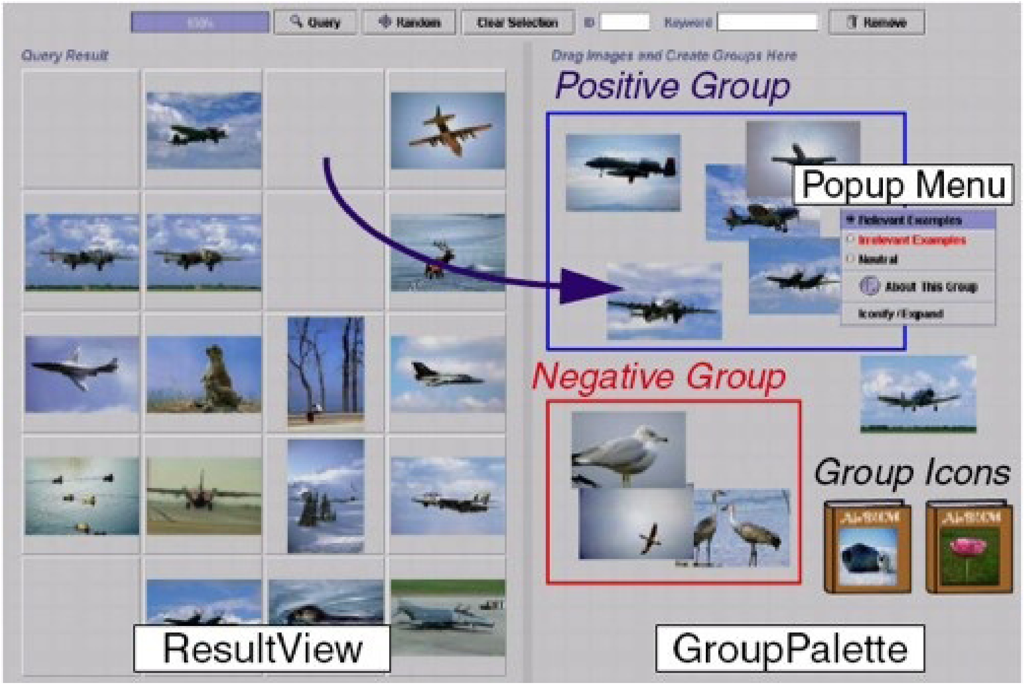

Figure 1
ImageGrouper interface

a structured query language based on conceptual graph or derived from ontology syntax. A structured query allows users to express and describe the spatial organization of objects in whose images the user is interested. This kind of interfaces is associated with semantic approaches (Lux et al., 2003) as well as to emergent statistical methods of indexing (Town and Sinclair, 2004).

It is difficult to evaluate how successful image retrieval systems are in terms of effectiveness, efficiency, and flexibility. Of course there are the notions of precision (the ratio of relevant images to the total number of images retrieved) and recall (the percentage of relevant images among all possible relevant images). Many articles about systems give figures about precision and recall. Most of them are good but hard to verify. In order to be able to compare the effectiveness of these different approaches, some test bases associated with sets of predefined request-answers must be built. This kind of experiences exists for text retrieval systems in the TREC network (http://trec.nist.gov: May 2007) but in the case of image retrieval the needs are so varied (precise or broad, analogical or identical, subjective or real) that the definition of an experimentation support is very complex. However, one should note that the INEX network proposes such an approach for the multi-media retrieval systems using XML (Fuhr N. et al., 2005).

Discussion

Image indexing and retrieval represents one of the remaining challenges in information systems. The problems of indexing and retrieval are compounded by the specific demands of the application domain (architecture and building). This includes the various formats of architectural images (pixel and vector-based, 2D and 3D, static and dynamic) and the different points of entry in query formulation. On the positive side, developments in indexing and retrieval stress that effective and efficient methods rely increasingly on domain analysis, not only technology and knowledge transfers from other areas. This supports the correlation of domain knowledge, from architectural perception and representation to codifications of concepts and entities, to indexing methods and approaches to retrieval.

The correlation of domain requirements with indexing methods and approaches to retrieval returns a distinction of *structural* and *semantic* subjects. The structural subjects refer to fundamental choices and problems, relating to e.g. the size and degree of homogeneity of an image collection, as well as to implementation issues, such as the colour, texture and shape parameters in CBIR – issues related to the analysis depictive representations. The semantic ones link the meaning and segmentation of architectural representations to CBIR and semantic approaches to indexing (e.g. the usability of design elements as visual descriptors and the degree of relationship between authority lists, information standardization and shared ontologies in architectural design).

In terms of usability, it is evident that automation is of paramount importance to most indexing situation, also with respect to reliability and objectivity. In retrieval usability relates strongly to user-friendliness, a probably antiquated and vague concept, which nevertheless expresses the problems of the user who has to balance not only between subjective interpretation and the objective structure of representations and terminologies but also between the abstraction of a query and the specificity of required information. The abstraction levels of architectural representations provide useful insights into these issues, as well as ready solutions to interface and implementation problems.

References

Aygen, Z. and Flemming, U.: 1998, Classification of precedents, CAADRIA '98, CAADRIA, Osaka, pp. 435-444.

Bignon, J.-C, Halin, G. and Nakapan, W.: 2000, Building product information search by images, in H. Timmermans (ed), Fifth Design and Decision Support

Systems in Architecture and Urban Planning, TU Eindhoven, Nijkerk.
Brown, F. E. and Steadman, J. P.: 1986, A computerised database of contemporary house plans, Environment and Planning b: Planning and Design, 18, pp. 405-414.
Del Bimbo, A.: 1999, Visual Information Retrieval, Morgan Kaufmann, San Francisco, USA.
Fuhr N., Lalmas M, Malik S. and Z, S.: 2005 INEX 2005 workshop, in Springer (ed), INEX, Dagstuhl, Germany.
Gross, M. D.: 1995, Indexing visual databases of designs with diagrams, in A. Koutamanis, H. Timmermans and I. Vermeulen (eds), Visual databases in architecture. Recent advances in design and decision making, Avebury, Aldershot.
Halin, G., Créhange, M. and Kerekes, P.: 1990 Machine learning and vectorial matching for an image retrieval model, in ACM (ed), SIGIR '90. 13th International Conference on Research and Development in Information Retrieval, Bruxelles, Belgique.
Kacher, S., Bignon, J.-C., Halin, G. and Duffing, G.: 2002, The content-based image retrieval as an assistance tool to the architectural design, in H. Timmermans (ed), Sixth Design and Decision Support Systems in Architecture and Urban Planning, TU Eindhoven, Avegoor.
Kosslyn, S. M.: 1994, Image and brain: the resolution of the imagery debate, MIT Press, Cambridge, Massachusetts.
Koutamanis, A., Halin, G. and Kvan, T.: 2007, Information standardization from a design perspective, CAADRIA 2007, CAADRIA, Nanjing.
Koutamanis, A.: 2000, Representations from generative systems, in J. S. Gero (ed), Artificial Intelligence in Design '00, Kluwer, Dordrecht.
Koutamanis, A.: 2006, Recognizing architectural representations, in V. Bourdakis and D. Charitos (eds), Communicating space(s). eCAADe, Volos, pp. 384-391.
Kvan, T.: 2004, Modes of representation, Architectural Review Australia, AR090, pp. 82-83.
Li, J. and Wang, J. Z.: 2003, Automatic linguistic indexing of pictures by a statistical modeling approach, EEE Transactions on Pattern Analysis and Machine Intelligence, 25(9), pp. 1075-1088.
Lux, M., Becker, J. and Krottmaier, H.: 2003 Caliph & Emir: Semantic Annotation and Retrieval in Personal Digital Photo Libraries, CAiSE '03, 15th Conference on Advanced Information Systems Engineering, Velden, Austria.
Mezaris, V., Kompatsiaris, I. and Strintzis, M. G.: 2003 An Ontology Approach to Object-based Image Retrieval, in IEEE (ed), Image Processing(ICIP03), Barcelona.
Nakapan, W., Halin, G, Bignon, J. and Wagner, M.: 2004, Extraction of building product image from the Web, International Journal of Intelligent Systems, 19(1-2), pp. 65 - 78.
Nakazato, M., Manola, L. and Huang, T. S.: 2003, ImageGrouper: a group-oriented user interface for content-based image retrieval and digital image arrangement, Visual Languages and Computing, 14(4), pp. 363-386.
Polovina, S. and Heaton, J.: 1992, An Introduction to Conceptual Graphs, AI Expert, pp. 36-43.
Porter, V. and Thornes, R.: 1994, A guide to the description of architectural drawings, G.K. Hall, New York.
Town, C. P. and Sinclair, D. A.: 2004, Language-based Querying of Image Collections on the basis of an Extensible Ontology, International Journal of Image and Vision Computing, 22(3), pp. 251-267.
Veltkamp, R. C. and Tanase, M.: 2002 Content-Based Image Retrieval Systems: A Survey, Department of Computer Science, Utretch University.
Vleugels, J. and Veltkamp, R. C.: 2002, Efficient image retrieval through vantage objects, Pattern Recognition, 35(1), pp. 69-80.
Yessios, C. I.: 1987, The computability of void architectural modeling, in Y. E. Kaluda (ed), Computability of design, Wiley, New York.