AUTOMATED INTEGRATION OF CONSTRUCTION IMAGES IN MODEL BASED SYSTEMS

Ioannis K. Brilakis [1] and Lucio Soibelman [2]

[1] Assistant Professor, Department of Civil and Environmental Engineering, University of Michigan, Ann Arbor, USA
[2] Associate Professor, Department of Civil and Environmental Engineering, Carnegie Mellon University, Pittsburgh, USA
Correspond to lastname@umich.edu (brilakis)

ABSTRACT : In the modern, distributed and dynamic construction environment it is important to exchange information from different sources and in different data formats in order to improve the processes supported by these systems. Previous research has demonstrated that (i) a significant percentage of construction data is stored in semi-structured or unstructured data formats (ii) locating and identifying such data that are needed for the important decision making processes is a very hard and time-consuming task. In this paper, an automated methodology for the classification and retrieval of construction images in AEC/FM model based systems will be presented. Specifically, a combination of techniques from the areas of image processing, computer vision, and content-based image retrieval have been deployed to develop a method that can retrieve related construction site image data from components of a project model.

Key words: images, materials, model based systems, image processing

1. INTRODUCTION

In the last decade the amount of data and especially image and video data has experienced an unparalleled growth due to technological developments in information technologies and digital storage as well as in the area of digital imaging. Lyman et al [1] in a study called "How much Information?" has shown that almost 375 petabytes (10^{15} bytes, or 787.5 billion photographs) of photographs are produced each year (almost 2 times all printed material) with a yearly growth rate of 5%, which is the highest growth rate among data types produced. At the same time, although the number of film-based photographs is declining at a 10% yearly rate, there has been a dramatic growth in the creation of images using digital cameras. For instance, 27.5 million digital cameras were purchased worldwide in 2002 which represents a yearly increase of 100% raising the amount of digital cameras to 53 million within a year compared to 63 million analog cameras that existed in the same year. Over 60 million digital cameras were sold in 2004, and more than 70 million are expected to be shipped in 2005 [2].

The same trend can also be observed in the AEC/FM industry and specifically in construction where traditional cameras are quickly replaced with their digital counterparts due to the latter's affordability, high resolution imaging, industry-wide acceptance and ease of use. The acquired images are stored in digital format and then uploaded to a central database that represents the site photographs log. These images and other visual data are then used for visualization, progress monitoring, productivity measurements, and other construction management tasks.

For these uses it is common for owners to require contractors to take frequent pictures of their work while liability claims are forcing companies to keep a complete site photographs log of the evolving project as evidence of the "as-built" and the existing site conditions of the project site and its neighboring structures. Wood [3] demonstrated the importance of keeping good image records for litigation by presenting several cases where judges overruled experts that failed to consider photograph evidences preferring to accept uncritically the contractor's untested accounts. All these facts demonstrate that images have evolved into a significant and irreplaceable part of the project documentation and justify the reasons for the ever-growing imaging information acquisition rate of the AEC/FM industry.

However, from a study of 5 large scale construction projects (data provided by several companies) we found that construction companies tend to store thousands of images without following any standardized indexing protocols. Thus, as the number of images in each project database grows larger, the task of locating and retrieving a single image (or a group of) becomes harder, more tedious and more error-prone. This effect is magnified by the amount of time needed to identify and record the entire content of each construction site image. In reality, engineers are accustomed to take each snapshot for a specific purpose, and therefore images are usually poorly classified and serve only a limited number of predetermined tasks. For example, a site engineer took a snapshot (Fig. 1) of the domestic water mains of a structure and classified it accordingly, thus neglecting the beams, columns, slabs and other neighboring structures that were also depicted in that snapshot. Since the success of every construction project is linked to the ability of accessing project information in a fast and efficient manner, models for automated classification and retrieval of construction site image data are needed.

Figure 1. Image labeled “domesticwatermains.jpg”

This paper reports the latest developments of the authors’ on-going research efforts in construction site image retrieval. The available models for image classification and retrieval in construction and other industries are initially outlined, and their benefits and limitations are explained. Following that, a novel automated image retrieval model is presented. This model can retrieve images from related objects in project models or construction databases using material, spatial and temporal information. The results demonstrate that this approach can effectively overcome the limitations of existing models and provide the flexibility of retrieving images either from object oriented project models or construction databases.

2. IMAGE CLASSIFICATION AND RETRIEVAL IN THE CONSTRUCTION INDUSTRY

Since the beginning of the recent information explosion era when construction companies gradually started to replace traditional cameras and image file cabinets with digital cameras and electronic image databases, several research efforts began addressing the image retrieval concerns described above. A prototype relational database using Microsoft Access was developed [4] where the engineer could manually link construction multimedia (audio/video) with other construction items of a database. This work stressed the importance of indexing construction multimedia files of all types and provided a solid model for resolving this issue. However it did not provide a solution that would help avoid the time consuming and tedious manual indexing that was still required while the problem of retrieving images of materials in inventory or temporary facilities that are not part of the “final product” was not solved.

Livengood [5] studied the significance of keeping a well maintained photograph database from a litigation perspective and reached similar conclusions; i) Construction projects are overwhelmed by photographs, ii) Photographs form perhaps the most effective mode of documenting contemporaneous events for use in claims and disputes, iii) A picture is truly “worth a thousand words”, iv) Participants in the construction process can make much more effective use of digital photography as a risk management tool

Based on these conclusions, Livengood [5] suggested a commercial manual photograph management software called “LYNX photo documentation” that is based on the principles of manual classification and promotes multi-modal searching. This system can also interface with Primavera and Meridian products.

Thesauri have also been suggested [6] as an approach that could assist in the image indexing processes. Such an approach, although it could automate image indexing to an extent, would still require the user to manually label each image using specific labeling standards and would therefore be semi-automated. Moreover, the search criteria that a construction manager could use when looking for project images would be still limited by the predetermined classification of the site engineer that usually targets only specific information.

Recently, we developed a construction site image retrieval model [7] based on an existing content based image retrieval (CBIR) research. This model utilizes blind relevance feedback techniques from the text mining area to automate a relevance feedback CBIR approach. After careful evaluation, this model was shown to boost the results of conventional CBIR models (when applied to construction site image databases) by providing them with quality feedback and requiring minimal interaction or intervention from the user and his expertise. However, we later realized that the high frequency of similar images in construction databases reduced the usefulness of this approach. In other words, when using e.g. a structural image as an example to retrieve other similar images, this model would retrieve a significant percentage of the images in a typical construction database (similar structural images). Even though having to browse through only a fraction of the images is certainly beneficial, the actual amount of images can still be overwhelming when thousands of images are present.

3. IMAGE RETRIEVAL IN OTHER INDUSTRIES

In recent years Content Based Image Retrieval (CBIR) models have been a major topic of research and have been explored from many different points of view: from early heuristic-based feature weighting schemes [8] to recently-proposed optimal learning algorithms, probabilistic / Bayesian learning algorithms, boosting techniques, discri-minant-EM algorithms [9], biased discriminant algorithms [10], support vector machines [11] and other kernel-based learning machines.

However, the retrieval capability of the generic image retrieval tools is limited by their generic scope of operation.

Specifically, some of the related limitations are:

1) These tools make little use of the domain specific characteristics of the construction industry. The concept of matching whole images with whole images without utilizing any of the texture, color or structure – specific characteristics of the domain to enhance the quality of the results by focusing the retrieval capability on the aspects of the AEC/FM industry has severe impacts on the quality of the results.

2) Retrieval methods based on the generic focus of the CBIR concept can hardly differentiate among images within a narrow domain database [12] (e.g. images of structural elements). The narrow scope of the construction domain can provide significant advantages in the design and operation of a method created to serve a large but finite number of construction operations.

3) The ability of retrieving most relevant images is severely limited by the precision-oriented philosophy of the CBIR concept. When searching a generic database of millions of images it is often not useful to retrieve most relevant images since the relevant ones can be thousands or hundreds of thousands. For this reason, the generic CBIR search engines focus in retrieving a small sample of relevant results with as few non relevant images as possible.

4. MATERIAL BASED CLASSIFICATION OF CONSTRUCTION SITE IMAGES

This section presents our latest efforts in construction site image classification [13], which is the basis for the novel multi-modal image retrieval model that will be presented in the next section. The purpose is to assist the reader in understanding some of the main concepts used throughout the development of this continuing research.

Our previous investigation revealed that the initial CBIR approach had to be redesigned and modified in order to take advantage of the construction domain characteristics. These modifications were based on the need for:

1) Matching parts of each image instead of the entire content. In most construction site images only a part of each picture is related to the domain while the remaining parts are redundant, misleading and can possibly reduce the quality of the results. For this purpose, it is necessary to effectively crop the picture in order to isolate construction related items (pavement, concrete, steel, etc.) from picture background (sky, clouds, sun, etc.) or foreground (trees, birds, butterflies, cars, etc.).

2) Comparing images based on construction related content. Each relevant part of the picture needs to be identified with construction related terms and the comparison of images with other images or with objects in a model based system should not be performed at a low level (using color, texture, etc.). Instead, the comparison should be based on features such as construction materials, objects and attributes.

Overall, this material based classification method is comprised of 4 steps. In the first step, each image is decomposed into its basic features (color, texture, structure, etc) by applying a series of filters through averaging, convolution and other techniques. The image is then cropped into regions using clustering and the feature signatures of each cluster are computed. During the fourth step, the meaningful image clusters are identified and isolated from the rest by comparing each cluster signature with the feature signatures of materials in a database of material image samples called "knowledge base". The extracted information (construction materials found) are then used to classify each image accordingly. This method was tested on a collection of over a thousand images from several projects. The results showed that images can be successfully classified this way according to the construction materials visible within the image content.

5. IMAGE RETRIEVAL MODEL

After completing the image classification step, an image retrieval model was designed. This model was developed based on the need for:

1) An all-inclusive approach that can combine most available search criteria. Material, temporal and spatial criteria can be combined to limit the search space and answer queries in different ways. For example, searching with date & materials is important when monitoring the progress of an activity, while searching with location & materials is important when looking for evidence of faulty construction for litigation purposes.

2) Interface flexibility. Different companies use different information management interfaces like project databases, model-based systems etc. Developing a model that works with only one possible interface would severely limit its applicability. Instead, this model was designed to interface with either construction databases of any type or object-oriented, model-based systems.

3) Reducing the amount of user-intervention. The major objective of this research is to relieve the engineer from monotonous, laborious and time-consuming tasks that are not value-adding. For the purposes of this model, the goal is to provide a simple, user-friendly and easy to use retrieval model that reduces the time needed to retrieve construction site images.

It is important to note that this model is based on the assumption that some or all image attributes are readily available in the image repository (Fig. 2) through the use of existing methods. Materials, for example, can be automatically detected using the method described in section 4 [13]. Date and time data can be automatically assigned to each image file from the digital camera, and for location (2-dimensional (x,y) and 3-dimensional (x,y,z)) a number of positioning technologies can be used (Global Positioning Systems, Local Positioning Systems, etc.). There are also commercially available cameras with built-in GPS and altimeter capabilities.

The processes of this image retrieval model as shown in Fig. 2 start from a model based system or a construction database, where the user selects the object for which relevant imaging information is needed, and requests for images. In the model based system, the attributes of the selected object (materials, date of construction, location, etc) can be

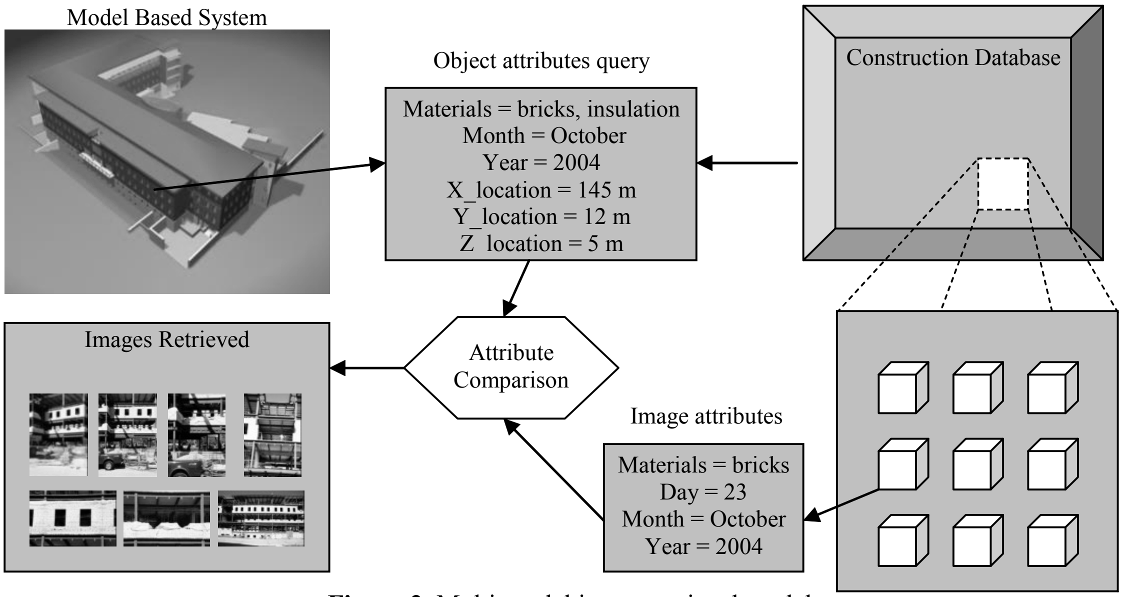

Figure 2. Multi-modal image retrieval model

automatically extracted using for example an IFC representation and used to formulate a search query, e.g. {materials = concrete, paint & month of construction = September & year of construction = 2004 & longitudinal location = 25m & latitudinal location = 5m & altitude = 7m}. Following that, the attributes of every image in the database are compared with the query's criteria in order to rank the available images according to their relevance. The results are then displayed on the screen.

The comparison method for each attribute depends on the attribute's type and the user's preferences. Materials, for example, are compared by matching the material names and using the amount of each material inside the picture content to further differentiate images with similar materials. Spatial and temporal data are continuous and are compared numerically (using the normalized Euclidean distance) and assist in limiting the search space for the material comparison. Temporal data are converted to a single numeric value before comparison by combining the time (ss/mm/hh) and date (dd/mm/yy) information. This way, the user can also search for images before or after certain dates, or images close to a certain date. The images are then sorted, based on the normalized Euclidean distance of the attribute vectors of the object and each image.

The result of this process is identical to the result of traditional search engines; a vector that contains all the images sorted according to the search query's criteria. The user is then responsible for browsing through the sorted images to identify and select the desired ones. This way, the user retains the flexibility of deciding which images are relevant in each case without having to browse through the entire image collection. The desired images are expected to most likely appear in the top of the sorted vector and so browsing though only these images is much faster and efficient.

From a user standpoint, this is a 2-step process, since when relevant images are requested, these images are retrieved and displayed immediately. The intermediate, hidden processes are real-time due the pre-assignment/pre-recognition of the attributes of each image. The sorted results can then be manually observed so as to extract the preferred images.

6. MODEL IMPLEMENTATION

In order to assess the validity of the model described above, a prototype implementation was developed that served as a test bed during the validation process. The prototype's architecture is comprised of three basic processes, the image objects formation, the image retrieval from a database and the image retrieval from a project model. The first process takes place when new images are entered into the company's image database. As the engineer uploads the new images the proposed methodology analyzes each image, crops it into regions and identifies the materials present with the help of the knowledge database. The image, material information and other available information are then assembled and stored in the form of an image object within the image database. This process is the most computationally intensive part of the model since it utilizes algorithms the scan the entire image several times, for example the fast Fourier transforms performed to compute the convolutions needed for the segmentation and texture transformations, the customized cropping / clustering / signature computing algorithm and the customized cluster/sample comparison algorithm. However the time cost to the engineer is negligent due to the complete automation of this process. This part is required since it populates the database of image objects, and works essentially as a pre-computation step for the following two processes.

The second process involves the image retrieval from a database and takes place when an engineer is searching for related images in a construction database. During this process the user-selected attributes of the image objects are compared with the corresponding attributes of other image objects in the database in order to retrieve a selection of images that share a number of similarities. The computational needs for each query are minimal since this process follows the process described above and thus the results can be displayed real time.

The third process involves the comparison among project model and image objects. Similar to the previous process, the objects' attributes are compared to define similarity among them. However, the major difference in terms of architecture is that the image database as well as the retrieval methodology is operated as a background process that is administered by the project model itself. In our implementation, Microsoft Visio operated as the front end of the system containing an IFC project model and calls to the developed prototype were made on demand from the user's selections. Specifically, when the user selects an IFC object and activates a custom macro by requesting for related images, this macro searches for the object's materials (IFCmaterial), location (IFCaxis2Placement) and date/time (IFCscheduleTimeControl). The information gathered is then used as a set of parameters that populate the search query. The computed distance of each image is then used to rank them in a results vector starting from the smallest distance. Following that, the prototype displays the images in a sequential manner, thus allowing the engineer to choose the relevant ones.

7. TESTING, RESULTS AND COMPARISON

The validation tests were conducted on an image collection of 1025 images. These images were grouped into 20 groups of related images based on their material information with an average size of 101.7 images/group. From these groups, the following seven materials provided a large enough data set to be statistically significant and thus chosen for further testing: Wood, earth, concrete, forms, rebar, steel, and paint. Following that, 7 material samples for each group were extracted from a separate subset of 30 images and were assembled to form the knowledge base. All images were then classified based on their material, date and location information, both manually and with the use of this prototype. The performance of the model was then assessed by comparing the results of the manual and the automated retrieval. Some sample comparisons can be seen in table 1.

Precision measures the percentage of relevant images retrieved over all images retrieved, and recall measures the percentage of relevant images retrieved over all relevant images. Scope represents the percentage of top-ranked images selected from the entire image collection. The precision at various recall points in Table 1 displays the performance of the proposed image retrieval model when compared with manual classification and retrieval. In each case, the precision reaches 100% when only the first few of the sorted images are considered, and decreases when the scope is increased. As expected, recall increases as more and more images are considered. Overall, it is important to note that precision stays high even at high recall levels, which shows robust performance for the proposed model.

The comparison of the results with those of manual classification and retrieval showed that the proposed model can automatically retrieve relevant construction site images with high precision and recall. The next step was to compare this approach in qualitative terms with the existing image retrieval approaches presented in section 2. In every case, the site engineer shoots images at the construction site, and the proceeds to upload them at the company storage database at regular time intervals. In manual classification and retrieval, the engineer labels each image according to its main purpose (usually a single object (or group of) in the image content) and must then create links of each image to every related object in the database. In thesauri based classification and retrieval, the engineer must use a certain thesauri standard to label each image explicitly, by providing enough information through the label so that the relevancies of each image with the objects in the database can be accurately computed. In the multi-modal image retrieval model, the engineer must acquire or create a collection of material samples needed for the knowledge base the first time of using the model. Every other time, the images must be uploaded to the company's database. For all cases of retrieval, every time that an image is needed in the future, the engineer can look for it through its related objects.

Table 1. Precision and recall at various scopes

Scope	1%	5%	10%	25%	50%	75%	100%
Material = Earth, 0 < Z < 3							
Precision	100%	89%	83%	72%	64%	49%	37%
Recall	1%	8%	15%	33%	59%	68%	68%
Material = Concrete, 10/10/2003 < date < 04/05/2004							
Precision	100%	86%	67%	56%	57%	64%	64%
Recall	1%	6%	10%	22%	45%	75%	100%
Material = Paint							
Precision	100%	62%	44%	41%	31%	21%	16%
Recall	2%	9%	13%	30%	46%	46%	46%

One very important difference of the proposed model with existing technologies is the amount of user intervention needed. Except the addition of spatial and temporal information to the image objects, all other processes in the multi modal image retrieval model are automated. Existing image retrieval methods rely on the user to define the links with related objects or input the keywords needed for this purpose.

Under certain and preset classification schemes, the existing construction site image retrieval methods provide accurate results. For example, if all images have been manually linked to corresponding objects in a database then the user can point to any database object and retrieve in real-time all relevant images without any redundancies. However, retrieval methods based on manual classification do not address the issue of how to index images dynamically, when unforeseen queries need to be addressed. (e.g. based on objects/components/categories not initially included in the database). Also, when looking (for example) for materials on site, equipment, temporary facilities or adjacent structures the existing classification schemes described in section 2 cannot provide any results, since there is no relevant object to link them with. One way to overcome this limitation is to avoid pre-classification schemes and retrieve images real-time based on more "flexible" criteria like visible objects and materials that are automatically extracted or available with temporal and spatial information. Therefore, when using the proposed multi-modal image retrieval model, this limitation is surmounted.

8. CONCLUSIONS

Every classification and retrieval approach has been developed with a certain use in mind and therefore is able to excel under specific circumstances. Manual indexing and retrieval is most accurate under certain pre-classification schemes at the expense of user-friendliness and flexibility. Thesauri-enhanced approaches, although more automated, are similarly tedious and error-prone and share the same disadvantages with the manual approach. Generic content based retrieval tools provide useful insights on how to automatically extract information from the image content so as to overcome the limitations of other methods. However their direct application in construction site image database systems and project models is severely limited by their generic design that is hardly applicable in narrow, domain-specific image databases. Multi-modal image retrieval combines the advantages of content based image retrieval methods and the domain-specific applicability of manual approaches into a user-friendly method that can produce high quality results.

The validation showed that the Multi-modal construction site image retrieval model can successfully answer object-attribute-generated image queries by comparing a wealth of common attributes of the images and the objects automatically. It retains and enhances the advantage of user-friendliness of the BRF approach while giving the opportunity to the engineer to retrieve images real time based on higher level domain specific concepts like materials, date and time instead of the low level concepts of color, texture and structure. Moreover this method addresses several of the issues and limitations of other methodologies discussed previously like taking advantage of the domain specific characteristics of construction and overcoming the problem specific deficiencies of the generic CBIR methods (e.g. low recall, focus on precision and wide domain databases, etc.).

REFERENCES

[1] Lyman, P. Varian, H.R. (2003) "How much Information?", Executive Summary, University of California at Berkeley, School of Information Science

[2] Kanellos, M. (2005) "Who will become the Intel of photography?" CNET News.com February 3, 2005

[3] Wood, S. "Why Keep Records", Proceedings of the Contractor and Construction Manager Issues 4th International Superconference London, UK, 19-20 May, 2005

[4] Abudayyeh, O.Y, "Audio/Visual Information in Construction Project Control," *Journal of Advances in Engineering Software,* Volume 28, Number 2, March, 1997

[5] Livengood, J. (2005) "Detailed Project Records: Sword and Shield" ", Proceedings of the Contractor and Construction Manager Issues 4th International Superconference London, UK, 19-20 May, 2005

[6] Kosovac, B., Froese, T. and Vanier, D. (2000) "Integrating Heterogeneous Data Representations in Model-Based AEC/FM Systems", Proceedings of International Conference on Construction Information Technology 2000 – CIB-W78, IABSE, EG-SEA-AI, Reykjavik, Iceland, 28-30 June, 2000, G. Gudnason (Ed.), Vol. 2, pp. 556-567

[7] Brilakis, I. and Soibelman, L., (2005) "Content-Based Search Engines for Construction Image Databases" *Automation in Construction*, Elsevier Science, Volume 14, Issue 4 , August 2005, Pages 537-550

[8] Ishikawa, Y., Subramanya, R. and Faloutsos, C. (1998) "MindReader: Query databases through multiple examples", in Proc. Of the 24th VLDB Conf. (New York)

[9] Rui, Y., Huang, T.S., Ortega, M. and Mehrotra, S. (1998) "Relevance Feedback: A Power Tool in Interactive Content-Based Image Retrieval", IEEE Tran on Circuits and Systems for Video Technology, Vol. 8, No. 5: 644-655

[10] Zhou, X.S. and Huang, T.S. (2001) "Comparing Discriminating Transformations and SVM for learning during Multimedia Retrieval", ACM Multimedia, Ottawa, Canada

[11] Hong, P., Tian, Q. and Huang, T.S. (2000) "Incorporate Support Vector Machines to Content-Based Image Retrieval with Relevance Feedback", IEEE Int. Conf. on Image Proc. (ICIP'2000), Vancouver, Canada

[12] Huijsmans, D. P. and Sebe, N. (2001) "Extended Performance Graphs for Cluster Retrieval" IEEE Conference on Computer Vision and Pattern Recognition (CVPR'01), vol I, pp. 26-31, Kauai, Hawaii, December, 2001

[13] Brilakis, I., Soibelman, L., and Shinagawa, Y. (2005) "Material-Based Construction Site Image Retrieval" *Journal of Computing in Civil Engineering*, American Society of Civil Engineers (ASCE), Vol. 19, No. 4, 2005