

Automatic Integration of Construction Images using Digital Imaging and Pattern Recognition Tools

Ioannis K. Brilakis

Visiting Scholar, Carnegie Mellon University, Pittsburgh, PA, USA

Lucio Soibelman

Associate Professor, Carnegie Mellon University, Pittsburgh, PA, USA

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 (iii) there is no automated method for associating visual data with their related project objects (iv) construction visual data and specifically images are a significant part of construction documentation with thousands stored in site photographs logs of large scale projects. Therefore, methods for managing such types of data and especially construction images are important for construction information management.

In this paper, an automated methodology for the retrieval, classification, and integration 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. This method is able to automatically classify, store, integrate and retrieve image data files in inter-organizational systems so as to allow their usage in project management related tasks like project monitoring and productivity measuring systems, processes visualization across time and inventory monitoring as well as evidence retrieval for litigation and other purposes.

Keywords Information retrieval, Images, Imaging techniques, Data retrieval, Information management

1. Introduction

The success of every construction project is linked to the ability of accessing project information in a fast and efficient manner. For example, when a project manager wants to find out when to order the steel beams and columns needed for a steel erection activity, the scheduling software is needed in order to access the scheduling data to find out when this activity will start. Later on, when the order must be placed, the estimating software is needed to find out the activity's cost, and the procurement software to find out who

the suppliers in the area are and what their prices are. This is one of the main reasons that model based systems are gaining ground during the last decade, since there has been a lot of research in methods for automatic integration of such types of data in project models, especially structured data (e.g. scheduling and estimating) that are kept in software-specific databases. The final goal is for example to be able to double-click on any object and retrieve any information related to that object like what materials it is made of, when it should be built and how much it will cost. However, so far no formalized method has been proposed for the automated integration of construction site pictures, and this is the problem that this research managed to solve.

The technological advancements in the area of digital imaging and the increasing popularity, availability and affordability of digital cameras and portable memory cards have given project managers the opportunity to keep detailed photographic logs of their projects. 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 takes a snapshot of the domestic water mains of a structure and classifies it accordingly, thus neglecting the beams, columns, slabs and other neighboring structures that are also depicted in that snapshot.

In this paper, a novel construction site image retrieval method is presented. This method was developed with the aforementioned considerations in mind and is based on pattern recognition techniques that originate from the areas of image and video processing and computer vision.

2. Research background

2.1 Data integration and file management in project models

During the last decade, the successful adoption of ISO STEP in many industries has influenced the development of product models such as Industry Foundation Classes (IFC) in construction. This development in object modeling in the construction industry was led by the International Alliance of Interoperability (IAI). Through its development of IFCs for model-based systems, IAI specified how objects that could occur in a constructed facility such as doors, walls, etc., and abstract concepts such as space, organization, process, etc., should be represented electronically. The IFCs are used to assemble a project model in a neutral computer language that describes building project objects and represents information requirements common throughout all industry processes (Froese et al, 1999). IFCs enable interoperability among Architectural-Engineering-Construction/Facility Management (AEC/FM) software applications. Software developers can use IFCs to create applications that use universal AEC/FM objects based on the IFC specification.

However, one of the most useful applications of IFCs in integrated model-based systems is the automation of file management. Project information is usually represented and exchanged as files. Files include design drawings, pictures, analysis calculations, bill of materials, schedules, contracts, etc. Managing and tracking project files is a crucial task in any AEC/FM project. The file management component in model-based systems is meant to implement the functionality of linking project files to the product model, thus enabling context-based access and management of these files.

2.2 Image processing and computer vision techniques for image retrieval

Almost all applications that make use of visual data in order to extract useful information are using methods from the areas of Digital Image and Video Processing, Computer Vision, Content Based Image and Video Retrieval and Statistics. These methods are based on the fact that images can be represented in a numerical three-dimensional matrix format (vertical (y), horizontal (x), color (RBGA)) and thus, mathematical concepts from algebra, statistics, geometry and chaos theory can be applied like signal processing filters (median, mean, Gaussian), mathematical transformations (Fourier, Wavelets), mean and variance, convolution, singular value decomposition and fractals.

Several computer vision and image and video processing methods can be used for visual data retrieval and content extraction. These methods are based on feature evaluation methods, or, in other words, visual data content is retrieved based on an image feature like color, texture, shape or structure. When color is used as the discriminant feature for retrieval, color histograms are used to determine similarity and dissimilarity among images. When texture is used as the discriminant feature for retrieval, several methods are applicable (Bovik, 2000), like Gaussian or wavelet smoothing (for noise removal), Fourier analysis (low pass filtering, phase reconstruction), wavelet decomposition, Laplacian and oriented pyramids and Gabor filters. When shape or structure is used as the discriminant feature for retrieval and content extraction methods like (Forsyth and Ponce, 2002) median filtering (for edge sharpening), clustering (K-means, graph-theoretic), imaging geometry (Shape from shading), Hough transforms (for edge detection), edge detectors (Canny, Sobel, Prewitt, Roberts, Laplacian of Gaussian, zero-cross, etc.), affine structure from motion (structure detection from video-streams), active contour models, and stereopsis (3D Shape reconstruction).

2.3 Multiple Feature Retrieval Methods

The three classic models in information retrieval are Boolean models, vector models, and probabilistic models. These models consider that each item (document, image, etc.) is described by a set of representative index terms. An index term can be simply a word whose semantics helps in remembering the document's main themes, or the color histogram of an image as described above. Given a set of index terms for an image, one can notice that not all terms (color, texture, structure, shape) are equally useful for describing the image contents. Thus, it should be clear that different index terms have varying relevance when used to describe picture contents. This effect is the main concept of the state-of-the-art multiple-feature Content Based Image Retrieval (CBIR) Methods and is captured through the assignment of numerical weights to each index term of a picture. In recent years 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 (Ishikawa et al., 1998) to recently-proposed optimal learning algorithms, probabilistic / Bayesian learning algorithms, boosting techniques, discriminant-EM algorithms, biased discriminant algorithms (Zhou and Huang, 2001), support vector machines (Hong et al., 2000) and other kernel-based learning machines.

However, commercial search engines are designed for a generic audience that frequently does not know what it wants or has great difficulty in properly formulating its request. Therefore, the recall factor (relevant images retrieved / all relevant images) is severely limited by the precision-oriented (relevant images retrieved / images retrieved) philosophy of these models. In other words, since CBIR tools are designed with huge databases in mind (millions of images), there is no benefit in retrieving all the relevant images (possibly thousands) since the user will not look beyond e.g. the first hundred images. Instead, these tools focus in removing non relevant images from the top of the list and give a sample of the relevant images to the user. On the contrary, in the AEC/FM domain (where searching for a specific image is common), it is often critical to retrieve all the relevant images even at the expense of increasing the number of irrelevant retrieved images.

3. Material Based Construction Site Image Retrieval

A detailed description of the proposed image integration method is presented in this section. The main differences with the existing generic image retrieval tools are: 1) Matching parts of each image instead of the entire content, 2) Comparing images based on construction related content only, 3) Utilizing spatial and temporal information (if available). Overall, the proposed solution is comprised of 6 steps:

Step 1: Analyzing image into basic features. Similar to the generic CBIR concept, the first step in this image retrieval approach is to isolate a number of image features by applying certain filters on the initial image. So far there is no optimal combination of these features in the literature since different research problems and different image content can be approached in different ways. Therefore, a number of these features must be tested in order to experimentally define their optimal combination. This can be achieved by selectively computing each feature for a number of images (that contain a single construction material each) and compare the results to identify the features that best describe each material.

Step 2: Dividing image into regions/clusters. As mentioned before, slicing each image into meaningful regions (in terms of construction) can be an important advantage to the retrieval method. There are many possible methods for detecting those regions, mostly windowing (overlapping and non-overlapping) or clustering oriented. However, since not all construction materials have a square or rectangular shape and their orientation as compared to the image plane is rarely parallel, windowing methods might pose the same problems as using the entire image content, since it can often be very hard to isolate a single material within a rectangular region. To avoid such issues clustering methods have been selected as more appropriate.

Furthermore, since construction materials tend to be monochromatic and uniformly textured, cluster creation should be based on segmentation related image transformations that can be derived by using the same techniques from the first step and a number of different segmentation filters like the Sobel, Canny, Marr-Hildreth and others, since such transformations detect and display most of the significant edges within the image content as well as their “strength”. Small sized filters (3x3) will also separate the noise pixels from the rest.

Step 3: Computing the signatures of each cluster. During or after the cluster formation step it is necessary to further compress the features of each image region into quantifiable, compact and accurate descriptors also known as *feature signatures*. In order to derive them, each feature has to be described in a way that accurately represents its content in all aspects and for this purpose statistics have been employed. Metrics like the mean, standard deviation, variance or the mode can accurately represent the pixel-values of each feature in a region, as long as those regions are relatively uniform (low variance).

Step 4: Identifying construction materials. Having sliced the given image into regions and computed their feature signatures we proceed to compare each signature with those in a knowledge database. This database is simply an image collection of material samples. Each sample is acquired from actual construction site images and must depict a single material (or part of). Its purpose is to give the method an example of what an already recognized material “looks like” for comparison purposes. The material image samples are then processed similarly to the given image, their features are extracted and their signatures are computed.

In order to identify the materials within the given image, each cluster signature is then compared to those in the knowledge database (normalized Euclidean distance), and, whenever a match is found, the cluster inherits the material name of the sample it was matched to.

Step 5: Assigning materials to original image. Altogether the given image along with the identified materials and other available information (e.g. spatial and temporal) form an image object. This object can

hold all available and extracted information about the picture's content, location and date acquired so as to assist in any future retrieval from other image objects or objects from model based systems.

Step 6: *Retrieving images*. After having assembled the image object the retrieval step can proceed by comparing each attribute of the object with the corresponding attributes of other image objects or model objects. In terms of materials, the comparison is performed real time by matching the material names. The resulting set of relevant images is then displayed to the user (Fig 1).

Fig. 1 Screenshots from the prototype implementation (Retrieval from MS Visio IFC Model)

4. Final Results

The final tests were conducted on an image collection of 876 images. These images were grouped into 20 sets of related images based on their material information with an average size of 101.7 images/group. The groups were defined based on the Masterformat standards as follows: Material Site Construction: wood (225 images), gravel (13 images), and earth (377 images), Concrete: Cast-in-Place concrete (279 images), scaffolds (105 images), forms (199 images), and rebar (267 images), Masonry: concrete blocks (17 images), Metal: steel (156 images), metal deck (27 images), metal framing (27 images), and metal studs (2 images), Thermal and Water Insulation: water (10 images), Finishes: drywall (7 images), paint (196 images), ceramic (40 images), Mechanical: pipes (64 images), exhaust (6 images), Electrical: conduits (11 images), smooth conduits (6 images). From these groups, the ones with statistically large number of images were 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. Each image was then cropped into regions and the clustering results are shown in table 1:

Table 1 Clustering Statistics

	Average	Variance	Std. Dev.	Min	Max	Max-Min
Avg. Pixels	3,479	4.6e+6	2,144	623	17,623	17,000
Pixels Found	77%	1.6%	13%	18%	96%	78%

The cluster quality was observed to be excellent (clusters did not penetrate the material borders) while the average number of pixels remained high. Clusters of such size (3,000 +) can accurately represent the material features within a confidence interval of more than 95%. The average percentage of pixels in large (>200) clusters (pixels found) is also more than satisfactory (77%) especially if we consider the amount of noise pixels in each image.

The clusters and their signatures were then compared with the knowledge database material samples. This comparison was performed using the entire feature set. The results were assembled into image objects and each object was used to run 7 queries, once for each material group. Fig 2 shows a summary of the method's performance for all tested material groups. It is important to note that the average precision stays high at high recall values and therefore retrieving most relevant images does not result in retrieving a lot of irrelevant images as well.

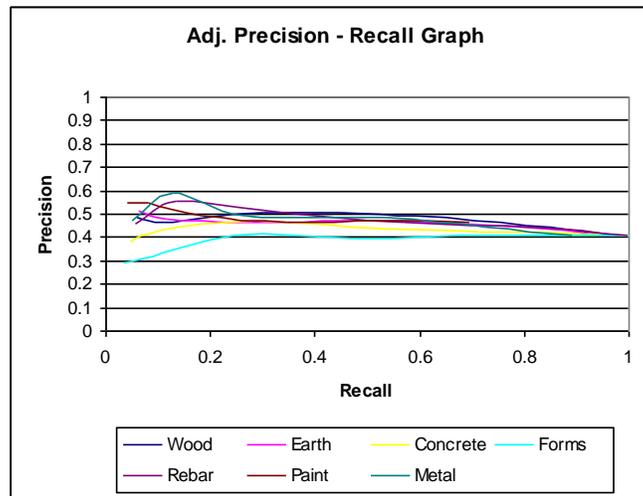


Fig 2 Precision Recall graph for seven material groups

5. Conclusions

This final evaluation has shown that the Material Based Construction Site Image Retrieval method can successfully answer material based image queries by pre-identifying the materials in each image and comparing material signatures instead of image signatures. 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 instead of the low level concepts of color, texture and structure. Moreover this method addresses all 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.).

6. References

- Bovik, A. (2000) "Handbook of Image and Video Processing". Academic Press, 1st edition (2000) ISBN:0-12-119790-5.
- Forsyth, D., and Ponce, J. (2002) "Computer Vision - A modern approach", Prentice Hall, 1st edition (August 14, 2002) ISBN: 0130851981.
- Froese, T., Fischer, M., Grobler, F., Ritzenthaler, J., Yu, K., Sutherland, S., Staub, S., Akinci, B., Akbas, R., Koo, B., Barron, A. and Kunz, J. (1999). "Industry Foundation Classes for Project Management – A trial implementation". Itcon Vol. 4, 1999.
- 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.
- Ishikawa, Y., Subramanya, R. and Faloutsos, C. (1998) "MindReader: Query databases through multiple examples", in Proc. Of the 24th VLDB Conf. (New York).
- Zhou, X.S. and Huang, T.S. (2001). "Comparing Discriminating Transformations and SVM for learning during Multimedia Retrieval", ACM Multimedia, Ottawa, Canada.