

SHAPE RECOGNITION OF LINEAR CONSTRUCTION ENTITIES FROM CONSTRUCTION SITE IMAGES

Ioannis K. Brilakis¹, and Lucio Soibelman²

ABSTRACT

The capability to automatically identify shapes, objects and materials from the image content through direct and indirect methodologies has enabled the development of several civil engineering related applications that assist in the design, construction and maintenance of construction projects. This capability is a product of the technological breakthroughs in the area of Image Processing that has allowed for the development of a large number of digital imaging applications in all industries.

In this paper, an automated and content based shape recognition model is presented. This model was devised to enhance the recognition capabilities of our existing material based image retrieval model. The shape recognition model is based on clustering techniques, and specifically those related with material and object segmentation. The model detects the borders of each previously detected material depicted in the image, examines its linearity (length/width ratio) and detects its orientation (horizontal/vertical). The results demonstrate the suitability of this model for construction site image retrieval purposes and reveal the capability of existing clustering technologies to accurately identify the shape of a wealth of materials from construction site images.

KEY WORDS

Shape, pattern recognition, construction site images, vision tools, objects.

¹ Assistant Professor, Civil and Env. Eng. Department, 2356 G.G. Brown Building, Univ. of Michigan, Ann Arbor, 2350 Hayward St MI 48109-2125, Phone +1 734/764-9420, FAX +1 734/764-4292, lastname@umich.edu (brilakis)

² Associate Professor, Civil and Env. Eng. Dept., 118N Porter Hall, Carnegie Mellon University, Pittsburgh, PA 15213, Phone +1 412/268-2952, FAX +1 412/268-7813, firstname@andrew.cmu.edu (lucio)

INTRODUCTION

Content-based (e.g. based on shape recognition) multimedia retrieval models have been applied to develop different types of systems, including: digital libraries, Web search engines, and specialized search engines (Brilakis and Soibelman, 2005b). The criteria and issues that have to be taken into account when building these types of applications are quite diverse. Regarding library applications, the criteria that the users adopt to judge document relevance are very important. In the Web, the collection of documents or images is substantial and encompasses a wide variety of subjects. Users have different backgrounds and preferences. Hence, commercial search engines must be designed for a generic audience that frequently does not know what he/she wants or has great difficulty in properly formulating his/her request. Search results usually return a large number of files with low relevance since the majority of the available files are not related with what the user is looking for.

Specialized retrieval applications, on the other hand, are developed with a particular application in mind. In such systems, the key problem is to retrieve (almost) all multimedia data that might be relevant without also retrieving a large number of non-relevant data. An example application of this type is construction multimedia information retrieval. Based on previous observations of the authors from project image databases (Brilakis and Soibelman, 2005f), construction site images are characterized by high volumes of very similar images with low topic variety as opposed to generic image databases. For example, in a typical building construction project, materials such as forms, wood, concrete, steel and earth can appear in over half of the available images and thus, a search based only on this criterion would, at best, reduce the image volume by half. The same often applies for temporal and spatial-based queries. In the same example, due to the limited space on the project site (as opposed to highway construction), the majority of images are shot from the same or a nearby location (at different stages of the project). The same low variability applies to the opposite examples on temporal queries since databases of repetitive projects or activities have low variability (e.g. in highway construction). For example, a paint subcontractor's image database contains images of various painting activities accumulated over time with often little or no visual information as to where exactly those images were acquired from. Similarly, asphalt-paving contractors' picture databases contain pictures of past paving activities that are highly similar and, without the extra temporal or spatial information, it is difficult to link them with their corresponding activities.

Such types of databases with low generality are often called "narrow-domain" databases. According to Huijsmans and Sebe (2001), how the performance of multimedia search methods degrades with rising generality greatly depends upon how similar or dissimilar the embedding items are with respect to the relevant items and how compact the clusters (groups) of relevant items are. In this light, it is important to emphasize the distinction of wide-domain versus narrow-domain databases: a wide domain database has a more uniformly filled feature space and since feature space in typical information retrieval systems is high-dimensional, performance will remain quite high even for a very large embedding (millions of items), especially when the clusters (groups of similar items) are quite compact; a narrow-domain database (e.g. a construction site image database as explained above) on the

other end of the spectrum is characterized by embedding items that are highly similar, and therefore, performance of generic information retrieval methods may degrade much more rapidly. Most civil engineering applications of content-based image or sound retrieval are more likely to be narrow- than wide domain databases; it is therefore crucial to develop features that remain distinctive in narrow-domain systems.

In other words, single feature (or a fixed combination of) retrieval in narrow domain databases such as construction multimedia databases is often not as effective as desired by the user. Multiple-feature (e.g. material+time+location) retrieval approaches are needed (Brilakis and Soibelman, 2005c). Also, even more distinctive features must be explored that will help take advantage of the unique visual characteristics of construction site embedding items and help in meaningfully differentiating construction images of high similarity (e.g. Fig 1).



Figure 1: Structural images. Embedding items are similar. Date and location are different

This paper presents a novel model for the automated recognition of two more distinctive image features; the linearity/non-linearity of the shapes of construction materials and their orientation in reference to the image plane. These novel features are a new addition to a multi-feature content based retrieval method that the authors have previously developed (Brilakis and Soibelman, 2005a). The detection of the new, shape-related features follows an earlier identification of the material content and is based on clustering (region-growing) techniques. The outcome is the recognition and subsequent distinction of linear and non-linear construction objects (e.g. concrete columns/beams vs. concrete walls) and well as the orientation of those objects detected as linear (e.g. steel columns vs. steel beams). The new model was implemented as an addition to the existing prototype from the authors' previous work and was tested separately. The results showed robust performance and validate the model's capability to further differentiate construction sites images.

MULTI-FEATURE RETRIEVAL OF CONSTRUCTION IMAGES

This section presents our latest efforts in construction site image classification (Brilakis and Soibelman, 2005a), which is the basis for the novel shape recognition model that will be presented in the following sections. The purpose is to assist the reader in understanding some of the main concepts used throughout the development of this continuing research.

As explained earlier, single feature retrieval in narrow domain databases (e.g. construction site image databases) is often ineffective in isolating the desired images. This fact, along with the continuing growth of the image volume in construction projects (Brilakis and Soibelman, 2005e), motivated the development of a multi-modal (multi-feature) image retrieval model that can retrieve images based on any combination of features. Specifically, this model was developed based on the need for:

- An all-inclusive approach that can combine most available search criteria. Three criteria (features) were initially incorporated into the model; date, location and materials. These 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.
- 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.
- Reducing the amount of user-intervention. The major objective of this research was to relieve the engineer from monotonous, laborious and time-consuming tasks that are not value-adding. For the purposes of this model, the goal was to provide a simple, user-friendly and easy to use retrieval model that reduces the time needed to retrieve construction site images.

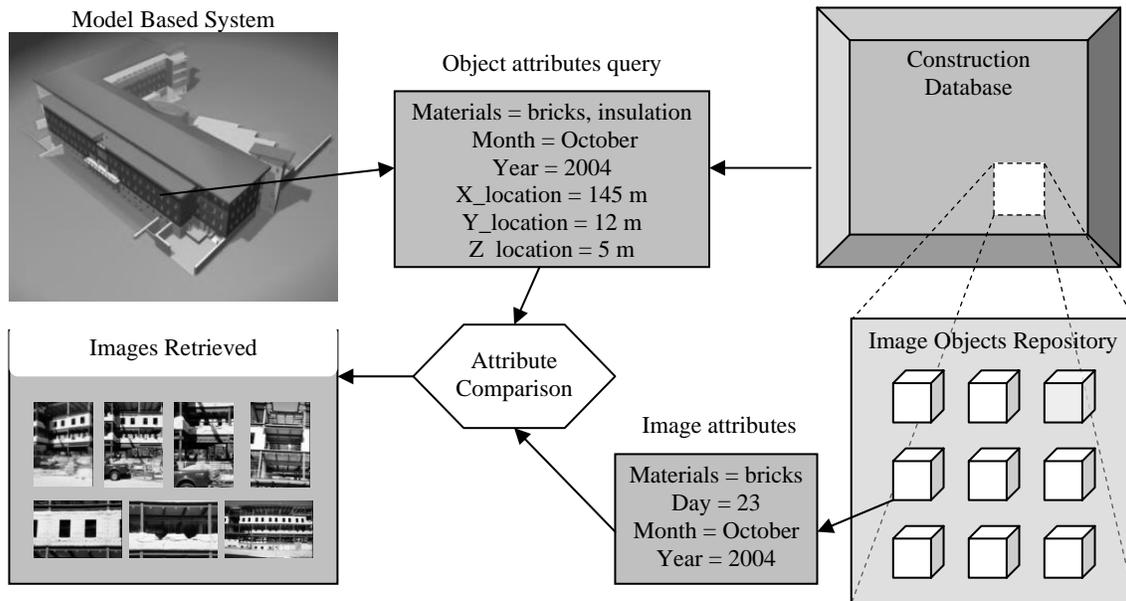


Fig. 2 Multi-modal image retrieval model

The processes of this 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 automatically extracted using for example, an IFC representation that can be 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.

SHAPE RECOGNITION

In order to extract the shape of objects in images, a method for representing shape is needed. One such representation, known as the medial axis, was introduced by Blum to describe the shapes of various biological structures (Blum, 1967). A refinement of the medial axis method, the multi-scale medial axis (MMA) technique, was introduced by Coggins (1992a-b). Both methods represent the shape of an object by applying a transform that skeletonizes the object. The skeleton (or medial axis) is defined as the set of all points in an object that has at least two points that are equidistant to the object's boundary. The medial axis provides a compact and convenient way of representing shape because regions become reduced to curves that follow the general shape of an object. This is useful in applications such as character recognition, in which shape information is more important than area or volume information. This work highlighted that i) Edges can be misleading because they often contain too much noise and yield information of poor quality; ii) Regions can be misleading because they often appear to be discontinuous; iii) The human vision system seems to preserve features and their related ambiguities at various spatial resolutions; competing hypotheses about an image can be preserved and resolved at the level of cognition rather than at the level of perception. As a result, Coggins reasoned that it is important to be able to interpret features on a continuous spatial-resolution scale.

Moses and Ullman (1992) constructed a theoretical argument for why non-model-based vision systems cannot in certain cases correctly recognize objects in a consistent manner. These authors offered a mathematical proof for their arguments based on a definition of consistent recognition functions. The main result of this proof is that because different objects can produce similar looking images or image features, it is not possible to distinguish these objects in those images (or with those features) without prior knowledge of how the images were formed.

For example, consider the cylinder and the truncated cone in Figure 3. The top and bottom as well as the left and right sides of the cylinder are parallel in view (a), and the left and right sides of the cone are not parallel. However, when viewed from orientation (b), the left and right sides of the cylinder are no longer parallel, making the cylinder indistinguishable from the cone. In order to reliably recognize these two shapes, a non-model-based vision system would have to be trained on all possible perspective transformations of these two objects. Furthermore, it would be difficult to take into account a

known viewing orientation to improve the discrimination capability of the non-model-based system. For this reason, this distinctive limitation of single, two-dimensional images that sets the boundaries of feature detection was considered throughout the development of our shape recognition model. Specifically, the proposed model was designed to operate in all cases, except those where three-dimensional information is needed to distinguish certain shapes, such as those in Fig. 3b.

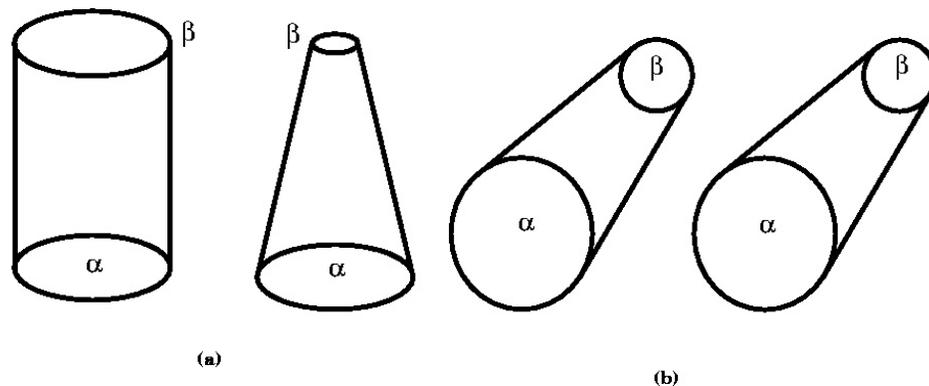


Figure 3: A cone and a cylinder (a) can have identical perspective transformations when viewed at a particular orientation (b) (Moses and Ullman, 1992).

THE PROPOSED CONSTRUCTION IMAGES SHAPE RECOGNITION MODEL

Shape recognition from construction images can assist in the development of several automated information extraction applications. In this particular case, a shape recognition model was developed to assist in better classifying and retrieving construction site images. As explained above, the multi-modal construction site image classification and retrieval model (Brilakis and Soibelman, 2005a) uses date, time, materials and manual schemes to classify and subsequently retrieve images from construction databases and model based systems. For several cases, such a distinction is sufficient. However, there are several materials that are frequently encountered in construction site images (e.g. concrete, steel, etc.) and, unless accurate spatial and temporal information are also available, image retrieval based on such materials could retrieve an overwhelming amount of pictures. In such circumstances, it is necessary to classify images in even smaller, more detailed groups based on additional characteristics that can be automatically recognized from the image content. Earth, for example, can be classified into the several different types of soil (Shin and Hryciw, 1999) while concrete and steel objects can be classified according to their shape (columns, beams, walls, etc.). The latter is what this shape recognition model can successfully recognize. Under this model, shape is represented as an additional feature in the multi-feature vector (cluster signature) used to mathematically describe each feature; the dimensions of each material region (cluster).

Similar to the approach of Blum (1967) and Coggins (1992) described previously, this model operates by skeletonizing the objects if such a skeleton exists. The difference lies in

the way of defining the skeleton, since the application is construction objects as opposed to biological substances. Objects in this case are presumed to be image areas of similar characteristics (e.g. similar color distribution, similar texture, or similar structure) with a certain degree of uniformity (since construction materials are often characterized by consistent colors, textures and structures). These image areas are selected using a flooding-based clustering algorithm (Brilakis et al, 2005) with high accuracy, and the materials that comprise each cluster are identified. Knowing the materials a priori increases the “intelligence” of the shape recognition, since in most cases, the majority of clusters are non-material clusters and so there is no need to recognize their shape.

Following that, the linearity and (if linear) orientation of the “object’s spine” of each cluster is evaluated. Both are determined by computing the maximum cluster dimension (MCD) and the maximum dimension along the perpendicular axis of MCD (PMCD) (Fig. 4). In the case of elongated polygonal objects, MCD is usually the line connecting the edges that are the furthest apart and can be considered as the object’s “spine”. Also, since PMCD can occur at any position along the perpendicular axis, every possible choice is evaluated. This step is essential since it also reduces the model’s vulnerability to noise (cluster topical inaccuracies) and fast, since it has a linear computation time ($O(n)$). These dimensions are then used to determine the linearity and orientation, under three assumptions:

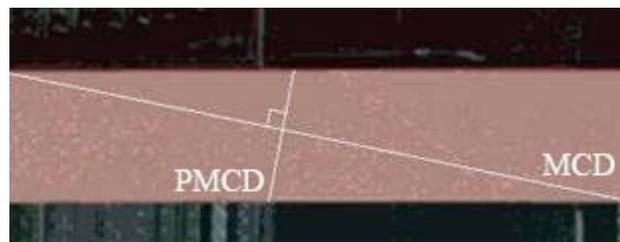


Figure 4: Steel cluster and measurements

- If MCD is significantly larger than PMCD, then the object is linear. This assumption is reasonable because construction objects have convex shapes in most cases.
- If the object is linear, then the tangent of the MCD edge points represents its direction on the image plane; the object’s “spine”. This assumption is stronger as the linearity of the object increases. Under different circumstances, the produced accuracy of the spine’s direction in reference to the image plane would be considered weak. For example, if in a rectangular shape MCD is 3 times PMCD, then the computed spine direction differs by 18.5 degrees from the actual spine direction. However in our case, we determined that construction beams and columns with a MCD over PMCD ratio of five or less are rare so the resulting maximum spine error is 11.3 degrees, which is more than enough for the purposes of simply detecting vertical (columns) and horizontal (beams) items and is therefore sufficient enough to draw reasonable conclusions.

- If the computed direction is within 45 degrees from the vertical/horizontal image axis then the linear object is a column/beam, respectively. In other words, it is assumed that the direction of the object is almost parallel to the image plane. This is an inevitable assumption, since as explained earlier (Moses and Ullman, 1992), a single image is not always enough to detect the desired information in a three dimensional space. Objects parallel to the focal length (e.g. the image of a beams' section) can convey little or no information about their other dimensions.

TESTING/RESULTS

The testing process for shape recognition was isolated from the testing process of other features (materials, time, location) to evaluate the performance of the shape recognition step. 103 images of steel and concrete beams and columns were isolated from the entire image collection provided that steel/concrete beams/columns:

- Actually exist and are of "significant" size within the image content. An empirical value of 200 pixels was used as the minimum "significant" size so as to crop image noise. The constraint induced by this minima is that objects that appear as minute within the image content are not considered.
- Are recognized as objects of concrete/steel by the remaining signature values. This way, shape recognition was solely responsible for the classification of this image subset.

Other than that, the performance criteria used for testing were precision (relevant over retrieved images ratio) and recall (relevant retrieved over total relevant images ratio). The results are depicted in Table 1.

Table 1: Shape recognition results

Material	Steel	Concrete
All images	103	103
Images retrieved	45	66
Non-relevant retrieved	4	12
Relevant not retrieved	1	6
Precision	91%	82%
Recall	98%	91%

The results presented in Table 1 demonstrate the capability of the proposed model in recognizing linear construction shapes, such as columns and beams made of concrete and steel, along with their orientation. The authors plan to test with even more examples in the near future, however the presented results are more than needed to validate this approach with statistical significance (95% confidence interval). It is interesting to note that the

performance of this model for steel items was higher. The authors believe that this is due to the lack of non-linear entities made of steel as opposed to the abundance of concrete slabs and walls (non-linear concrete entities). This effect also shows the robustness of this approach since the existence of non-linear items made of the same material does not significantly impact the recognition performance.

CONCLUSIONS

The shape recognition model presented in this paper was shown to robustly differentiate linear and non-linear construction entities and, provided that the entities are almost parallel to the image plane, to effectively detect their generic orientation (vertical/horizontal). This way, materials frequently seen in construction images can now be further differentiated into non-linear materials (e.g. concrete walls, concrete slabs, etc) and linear materials of certain directionality (e.g. steel columns, concrete beams, etc). This novel feature, when embedded within the multi-modal retrieval model, increases the retrieval flexibility and provides engineers with more search combinations to accommodate for more of their construction site image retrieval needs. For example, with this new feature addition, when searching for the date when a crack first appeared on a certain concrete beam, the query can be formulated as: i) material = concrete, ii) shape = linear, horizontal, and iii) location = x,y,z,direction. The addition of shape in this case will assign a low rank to most of the images depicting concrete slabs, walls and columns of that locality, and thus highlight images with concrete beams.

The findings of this work are also the first step in recognizing objects from construction site images. Aside from construction site image retrieval, construction objects recognition can potentially assist a large number of construction inspection and management applications such as productivity and progress monitoring, automated as-built/as-designed checks for deviation detection and others. These aspects of shape recognition are the next target of the authors, since there is great potential for major impacts on the construction industry by intelligently automating applications such as those stated above.

REFERENCES

- Blum, H. (1967) "A transformation for extracting new descriptors of shape" *Models for the Perception of Speech and Visual Form*, pages 362-380. MIT Press, Cambridge, Massachusetts, 1967
- Brilakis, I. and Soibelman, L., (2005a) "Multi-Modal Image Retrieval from Construction Databases and Model-Based Systems", *Journal of Construction Engineering and Management*, American Society of Civil Engineers, in review
- Brilakis, I. and Soibelman, L., (2005b) "Content-Based Search Engines for Construction Image Databases" *Automation in Construction*, Elsevier Science, Volume 14, Issue 4, August 2005, Pages 537-550
- Brilakis, I. and Soibelman, L., (2005c) "Automated Integration of Construction Images in Model Based Systems", *1st International Conference on Construction Engineering and Management*, 16-18 October 2005, Seoul, Korea

- Brilakis, I. and Soibelman, L., (2005d) "Automatic Integration of Construction Images using Digital Imaging and Pattern Recognition Tools", *3rd International Conference on Construction in the 21st Century*, 15-17 September 2005, Athens, Greece
- Brilakis, I. and Soibelman, L., (2005e) "Image processing techniques for identification of materials from construction site images", *International Conference on Computing in Civil Engineering*, 12-15 July 2005, Cancun, Mexico
- Brilakis, I. and Soibelman, L., (2005f) "Comparison of manual and user guided methodologies for the classification and retrieval of construction site images", *Construction Research Congress*, 5-7 April 2005, San Diego, California
- 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), Volume 19, Issue 4, October 2005
- Coggins, J.M. (1992) "Statistical investigations of multi-scale image structure" In *Proceedings of the Second International Conference on Visualization in Biomedical Computing*, Chapel Hill, North Carolina, October 1992
- 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
- Moses, Y. and Ullman, S. (1992) "Limitations of non model-based recognition" *Second European Conference on Computer Vision Proceedings*, pages 820-8, Berlin, Germany, May 1992. Springer-Verlag
- Shin, S. and Hryciw, R.D. (1999) "Wavelet Analysis of Soil Mass Images for Particle Size Determination" *Journal of Computing in Civil Engineering*, Vol. 18, No. 1, January 2004, pp. 19-27