Available online at www.sciencedirect.com
ScienceDirect
Advanced Engineering Informatics 20 (2006) 443–452
ADVANCED ENGINEERING INFORMATICS
www.elsevier.com/locate/aei

Construction site image retrieval based on material cluster recognition

Ioannis K. Brilakis [a,*], Lucio Soibelman [b], Yoshihisa Shinagawa [c]

[a] Department of Civil and Environmental Engineering, University of Michigan, 2340 G.G. Brown, 2350 Hayward St., Ann Arbor, MI 48109-2125, USA
[b] Carnegie Mellon University, Porter Hall 118N, Pittsburgh, PA 15213-3890, USA
[c] University of Illinois, 2021 Beckman, 405 N. Mathews Ave, Urbana, IL 61801, USA

Received 18 February 2006; accepted 2 March 2006

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 construction site image retrieval method is presented. This method is based on image retrieval techniques, and specifically those related with material and object identification and matches known material samples with material clusters within the image content. The results demonstrate the suitability of this method for construction site image retrieval purposes and reveal the capability of existing image processing technologies to accurately identify a wealth of materials from construction site images.

Keywords: Images; Construction; Image retrieval; Information retrieval; Materials; Clustering

1. Introduction

During the last decade, the applications of digital imaging in the civil engineering domain have significantly increased. The development of civil engineering tools that utilize image data to assist in the design, construction and maintenance of construction projects has been enabled by the capability to automatically identify shapes, objects and materials from the image content through direct (content-based) and indirect methodologies. For example, edge detection techniques are being utilized to detect the type and amount of surface cracks in pavements [1], motion segmentation to detect moving vehicles [2] and X-ray tomography to detect the internal structure of asphalt concrete [3] and to measure the shapes of aggregates [4]. Other examples include the usage of image color and intensity to assess fire-damaged mortar [5], image analysis to determine the strain distribution in geosynthetic tensile testing [6] and evaluate the fatigue of asphalt mixes [7], image velocimetry for flow diagnostics in fluid modeling [8], multi-resolution pattern classification of steel bridge coating [9], and electromagnetic image reconstruction for damage detection [10]. The work presented in this paper is focused on the application of digital imaging principles on sorting and retrieving construction site images.

The rate of digital imaging information acquisition is rapidly increasing. Digital cameras allow the engineers to take thousands of snapshots of a project throughout its life cycle and are gradually replacing traditional film cameras. During the conceptual phase, images of similar past projects and site conditions assist architects and engineers. During the construction phase, images of the work in progress assist in (a) remote progress monitoring, (b) identifying the differences between the design and the actual structure, (c) keeping evidence of the "as-built" (including temporary facilities and equipment), (d) resolving disputes (when utilized as evidence) and (e) video shots assisting in productivity measurements. During the maintenance phase,

* Corresponding author. Tel.: +1 734 764 8495; fax: +1 734 764 4292.
E-mail address: brilakis@umich.edu (I.K. Brilakis).

1474-0346/$ - see front matter
doi:10.1016/j.aei.2006.03.001

images assist in automatically detecting and recognizing damage (on pavements, walls, structural members, etc).

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 construction industry. However, from a study of five large scale construction projects (data provided from several companies) it was found that construction companies tend to store thousands of images without following any standardized indexing protocols, thus making the manual searching and retrieval a tedious and time-consuming effort. To overcome such limitations, the authors developed a novel material identification methodology that can effectively and automatically index construction site images according to their material information. This method removes the burden of manual classification from the user and allows for more flexible retrieval by not being tied to a certain indexing structure. The following section discusses a number of image retrieval techniques that have been investigated during the implementation of the material identification method.

It is important to note that the motivation for developing this method is twofold; effective and non-laborious indexing and retrieval of image data so as to address the concerns described above, and discovering the capabilities of cognitive tools for content extraction in images. The latter is especially significant, since accurate recognition and extraction of construction-related information (materials, objects, shapes, surface areas, etc) from the image content is the first step towards the development of more advanced and automated tools in the future. Such tools (e.g., real-time progress monitoring, automated productivity measurements and evidence detection and retrieval for litigation purposes) will significantly assist construction managers and other related disciplines.

2. Image retrieval based on content information extraction

Several computer vision and image and video processing methods can be used for visual data retrieval and content extraction. These methods utilize feature evaluation approaches, or, in other words, visual data content can be retrieved based on several features like color, texture, shape or structure.

When color is used as the discriminant feature for retrieval (Fig. 1), 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 [11], such as Gaussian or wavelet (for noise removal), Fourier analysis (low pass filtering, phase reconstruction) (Fig. 2), wavelet decomposition (Fig. 3 [12], Laplacian and oriented pyramids and Gabor filters. The advantage behind these tools is the explicit texture that results from applying these tools on images. For example, the texture on the surface of the wall in Figs. 2 and 3 is distinct from the texture of non-wall

Fig. 1. Color histograms detection. Wall surface detection based on color.

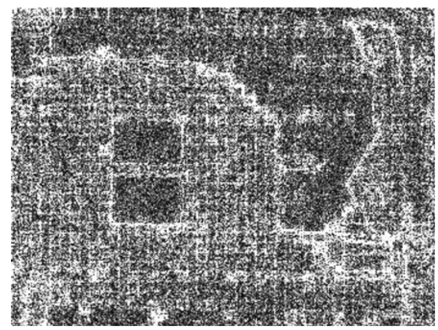

Fig. 2. Fourier analysis. Texture extraction from phase reconstruction.

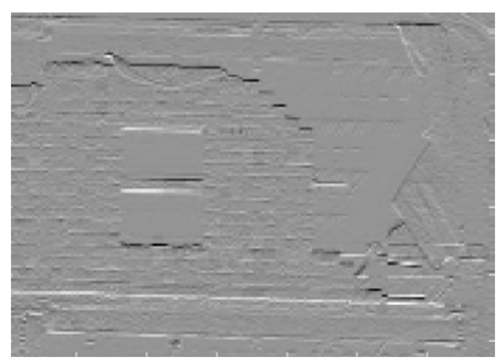

Fig. 3. Wavelet decomposition. Texture extraction from Level 2 Haar wavelets.

surfaces. From the human visual perspective these distinctions might not be very clear, but from the visual perspective of a computer, the distinction is mathematically sound [11].

When shape or structure is used as the discriminant feature for retrieval then content extraction methods like [13,14] median filtering (for edge sharpening), Hough transforms (for edge detection) [15], edge detectors (Fig. 4) [16], clustering (K-means, graph-theoretic) (Fig. 5), structure from motion (structure detection from video-streams), active contour models [17], and stereo-matching.

Overall, content-based image retrieval models have been applied to develop different types of systems, including: digital libraries, web search engines, and specialized search

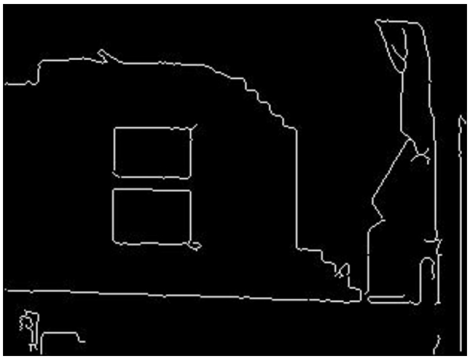

Fig. 4. Edge detection. Canny's edge detector.

Fig. 5. Clustering. K-means clustering.

engines. Specialized retrieval applications are developed with narrow domain databases in mind (e.g., for construction site images). In such systems, a common problem is to retrieve (almost) all relevant images, even if that results in retrieving a large number of non-relevant images. This problem is related with the increased significance of each image in such databases. For example, when trying to identify when a crack first appeared on a structural beam, the oldest image showing that crack can be the decisive proof for pinpoint who is at fault (supplier, contractor). On the other hand, generic retrieval applications are focused in retrieving relevant images with as few non-relevant as possible, even if only a small fraction of the relevant images is actually retrieved. For this reason, the retrieval capability of the generic image retrieval methods is limited by their generic scope of operation.

From our previous research efforts [18] it was concluded that 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 approaches (i.e., [19]) are most accurate under certain pre-classification schemes at the expense of user-friendliness and flexibility. Thesauri-enhanced approaches (i.e., [20]), although more automated, are similarly tedious and error-prone and shares the same disadvantages with the manual approach. The generic content based retrieval tools such as those described above 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.

3. Automatic materials identification from construction site images

According to what was described above, a novel material identification method for image retrieval was developed and is presented in this section. This method, as is going to be demonstrated at the validation section, can efficiently and reliably identify materials so as to index and retrieve construction site images automatically, thus removing the user from the tedious, error-prone and time-consuming classification process. Concepts from Content Based Image Retrieval (CBIR) approach were utilized for this purpose. However, a number of modifications were necessary in order to take advantage of the domain characteristics. These modifications were based on the need for:

(1) Matching parts of each image instead of the entire content.
(2) Comparing images based on higher-level, construction related content. CBIR algorithms match images based on low-level features like color and texture. By using those features to describe and identify construction materials it is then possible to retrieve images based on their material information. Thus, the user is able to utilize higher level concepts (materials in this case) to describe the query more accurately (e.g., find pictures with 80% concrete and 20% steel instead of find pictures with 35% red, 32% green, 33% blue) avoiding the retrieval of similar but unrelated images (e.g., when structural images are compared based on overall similarity characteristics they usually appear similar, so a structural image example will retrieve all structural images instead of just the desired ones).

3.1. Method description

A detailed description of the proposed material identification methodology is presented in this section. Overall, the proposed method is comprised of four 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 divided into regions using clustering and then the feature signatures of each cluster are computed. During the fourth step, the meaningful image clusters are isolated from the rest by comparing each cluster signature with the feature signatures of materials in a material knowledge base. The identified materials are then assigned to the original image as attributes, and finally those attributes are used to compute relevancies of images with other images in a construction image database.

3.1.1. Analyzing image into basic features

The first step in this material identification approach is to isolate a number of image features by applying certain filters on the original image. Examples of these features include the intensity, color histograms, several texture representations (filter banks, Gabor filters, wavelet coefficients, etc.) and others, as described earlier. So far, there is no optimal combination of these features provided by available literature since different research problems and different image content can be approached in different ways. Therefore, the following well established features were utilized: intensity, color distribution and responses to filter banks [21]. Specifically, 10 features were used; intensity (Eq. (1)) was computed as the average of the three color channels. Subsequently and in order to extract the intensity value from each color, the red, green and blue color channels were divided by the intensity to reflect the percentage contribution of each color to the total (Eq. (2)).

$$\text{Intensity }(x,y) = \frac{\text{Red }(x,y) + \text{Green }(x,y) + \text{Blue }(x,y)}{3} \tag{1}$$

$$\text{NRed }(x,y) = \frac{\text{Red }(x,y)}{3{*}\text{Intensity }(x,y)} \tag{2}$$

The remaining six features were texture related and measured the spottiness, and directionality of each texture. These features are computed by convolving the image with a series of filters (also known as filter kernels). The resulting output is an image of the same size that depicts the response of the original image to the kernel that it was convolved with. According to Forsyth and Ponce [13] the advantage of transforming the image to the new basis is that this process makes the local structure of the image clear. This is because there is a strong response when the image pattern in a neighborhood looks similar to the filter kernel and a weak response when it doesn't. Moreover, since there is no canonical answer in deciding which filters to choose, the filters used were based on the analogy with the human visual cortex based the work of Malik and Perona [21]. According to them, at least one spot filter and a collection of oriented bar filters at different orientations scales and phases should be used. The phase of the bar refers to the phase of a cross section perpendicular to the bar, thought of as a sinusoid (i.e., if the cross section passes through zero at the origin then the phase is 0°). One way to obtain these filters [21] is to form a weighted difference of Gaussian filters at different scales, and so the filters designed for this research are:

(1) A spot (Fig. 6), given by the weighted sum of three concentric, symmetric Gaussians, with weights 1, −2 and 1, and corresponding sigmas 0.62, 1, 1.6.
(2) Another spot, given by the weighted sum of two concentric, symmetric Gaussians, with weights 1 and −1, and corresponding sigmas 0.71, 1.14.

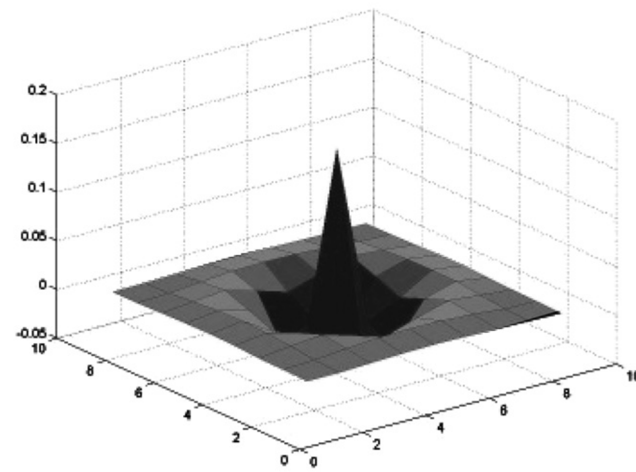

Fig. 6. Spot filter – weights: 1, −2, 1; sigmas: 0.62, 1, 1.6.

(3) A series of oriented bars (Figs. 7 and 8), consisting of a weighted sum of three oriented Gaussians which are offset with respect to one another. There are four

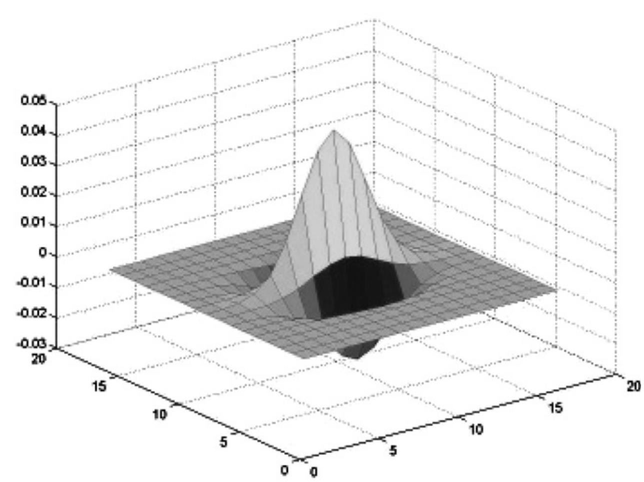

Fig. 7. Horizontal line filter – weights: 1, −2, 1; sigmas: $\sigma_x = 2$, $\sigma_y = 1$.

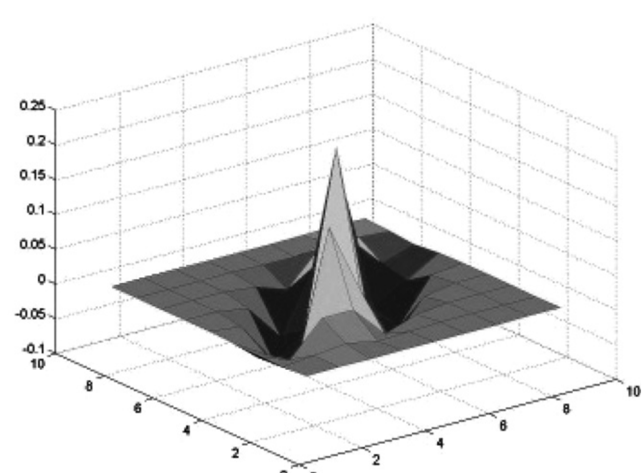

Fig. 8. Diagonal line filter – weights: 1, −2, 1; sigmas: $\sigma_x = 2$, $\sigma_y = 1$.

versions of these bars; each is a rotated version of the horizontal bar. The Gaussians in the horizontal bar have weights $-1, 2, -1$. They have different sigma in the x and in the y directions; the σ_x values are all 2 and the σ_y values are all 1. The centers are offset along the y axis, lying at (0,1), (0,0), (0,−1).

The same concept of convolving the image with a filter is used for segmentation purposes (cropping "similar" pixel areas like materials). The filters in this case were the gradient based edge detector kernels that were utilized for detecting the edges in the picture. In a gradient based edge detector, an estimate of the gradient magnitude is computed and used to determine the position of the edge points. In our case, three well established and commonly used filters were applied and their results were compared:

(1) A gradient filter, applied as a one dimensional filter to each of the two image dimensions.

$$[0, 1, 0, -1, 0] \tag{3}$$

(2) The Canny (or Normal Derivative of Gaussian with sigma <1) filter, applied similarly to the Sobel filter.

$$[0.00061811, 0.00578917, 0.00000000, -0.00578917, -0.00061811] \tag{4}$$

(3) The Marr–Hildreth (or Laplacian of Gaussian) filter, applied as a 5×5 two dimensional filter as shown in Fig. 9.

$$(\nabla^2 G)(x, y) = \frac{\partial^2 G}{\partial x^2} + \frac{\partial^2 G}{\partial y^2} \quad \text{(The Laplacian operator,[22])} \tag{5}$$

Convolving such kernels with thousands of construction images (of several millions of pixels each) posed a significant computational burden. Specifically, initial testing with direct convolution showed an average of 15 min processing time for each picture. To overcome this limitation, the convolution of each kernel with the image was performed using a fundamental property of the Fourier transform (FT) that substitutes the computationally expensive convolution process with a simple element by element multiplication of the FT of the image and of the kernel followed by an inverse FT of the result. The process used to perform each convolution using the FFTW library is shown in Fig. 10.

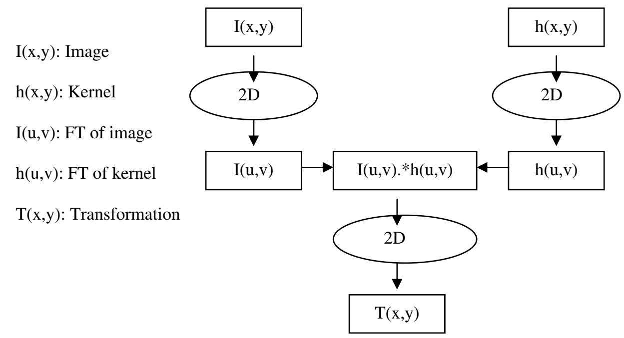

Fig. 10. Convolving a two dimensional image with a two dimensional filter using fast Fourier transforms.

$$\sim [I(x, y) \otimes h(x, y)] = I(u, v)(.^*)h(u, v) \quad ([23]) \tag{6}$$

After each convolution the values in the resulting transformation were normalized within the zero to one range and "folded" in half in order to increase the spread of values since the direction of the gradient (both for texture and for segmentation) is not important. In other words, for edge highlighting purposes, approaching an edge from the left or the right is of no importance. This way, the response of the smooth image regions is dark and the response of the edges is light (see Fig. 11).

3.1.2. Dividing image into regions/clusters

There are many possible methods for detecting single-material regions, mostly windowing (overlapping and non-overlapping) or clustering oriented. However, since not all construction materials have a square or rectangular

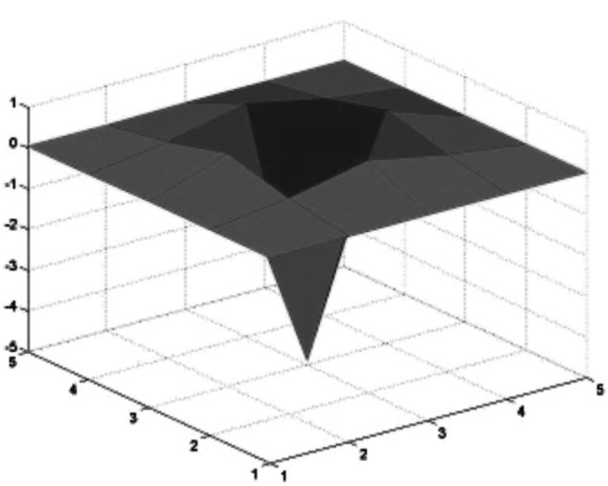

Fig. 9. Standard 5×5 Marr–Hildreth filter.

Fig. 11. Sobel filter applied on building facade (edges in white).

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 being more appropriate.

In order to effectively identify materials the clustering method of choice should result in clusters that contain all the image pixels and that do not violate the material borders significantly. Cropping a certain material into more than one region, although not desired, does not pose any significant problems to the overall retrieval approach. For this purpose, bottom-up clustering methods were selected as being the most appropriate.

During implementation, a customized clustering algorithm was developed that computes the signatures of each cluster while identifying them. Each pixel is visited only once, and a "flooding" process is initialized from it towards its neighbors. Every "similar" pixel is added to the cluster and marked so as not to be processed again later. Initially this process was designed to work recursively, however the extensive recursion levels (cluster size can exceed 3000 pixels) would result in stack overflows, so an alternative approach using stacks was implemented.

Furthermore, since construction materials tend to be monochromatic and uniformly textured, cluster creation should be based on edge detection related image transformations that can be derived by using the same techniques from the first step (analyzing the image) and a number of different edge detection filters that detect and display most of the significant edges within the image content as well as their "strength" that assists in the differentiation of edges so that weak edges (e.g., within a material texture like a brick wall) can be separated from the strong edges (e.g., the edge between a brick wall and a steel column). Furthermore, by using small sized filters (e.g., 3×3) the image noise pixels can be separated from the rest. Fig. 11 shows an edge detection example using the Sobel filter. Strong edges are marked in white, while weak edges are marked in grey.

For implementation purposes, gradient based edge detector kernels were utilized for detecting the edges in the picture and assist in the clustering step. In our case, three well known filters were tested with equally good results: Sobel, Canny and Marr–Hildreth.

After each convolution the values in the resulting transformation were normalized within the zero to one range and "folded" in half in order to increase the spread of values since the direction of the gradient (both for texture and for segmentation) is not important (approaching an edge from the left or the right is of no importance). This way, the response of the smooth image regions is dark and the response of the edges is light (see Fig. 11).

However, the most important concept of this step is the way of defining similarity among pixels. As mentioned before the segmentation transformations computed in the previous step define this comparison effectively. When an edge is met, the different response (white) prohibits the cluster to grow beyond the borders. However, not all edges are strong and as a result clusters would often find a "hole" among the edge perimeter to grow outside. This drawback was corrected as follows: whenever the gradient change was smooth and the resulting edge weak the algorithm compares each pixel intensity value with the pixel from where the cluster originates. In this way, as soon as the cluster would "escape" and start growing outside the desired region of values the absolute difference in intensity values would prevent it from growing further in that area.

3.1.3. 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 (Fig. 12). These signatures are the necessary data structures for the comparison with the material signatures from a knowledge database as described below. 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). This is a strong assumption in this case since each "useful" cluster is expected to contain pixels from a single material. For this implementation and as mentioned in the first step, ten features were utilized (intensity, three colors, six textures) and their mean and variance was computed during the clustering step. The result was a 20 dimensional signature for each cluster.

3.1.4. Identifying construction materials

Having cropped the given image into regions and computed their feature signatures, each signature is then compared with the material signatures in a knowledge database (Fig. 13, Table 1). 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 it). The purpose is to give the method an example of what a pre-classified material "looks like" for comparison purposes. After the labels are provided, the material image samples are added and they are processed similarly to the given image. In other words, their features are extracted and their signatures are computed and stored for future comparisons.

A single image sample for each material is enough to detect instances of that material. Additionally, more samples can be added to increase the detection precision. When new samples are introduced, an automated comparison

62	33.5	32.5	…	0.58	0.003	0.001	…
Means				Variances			

Fig. 12. Sample signature.

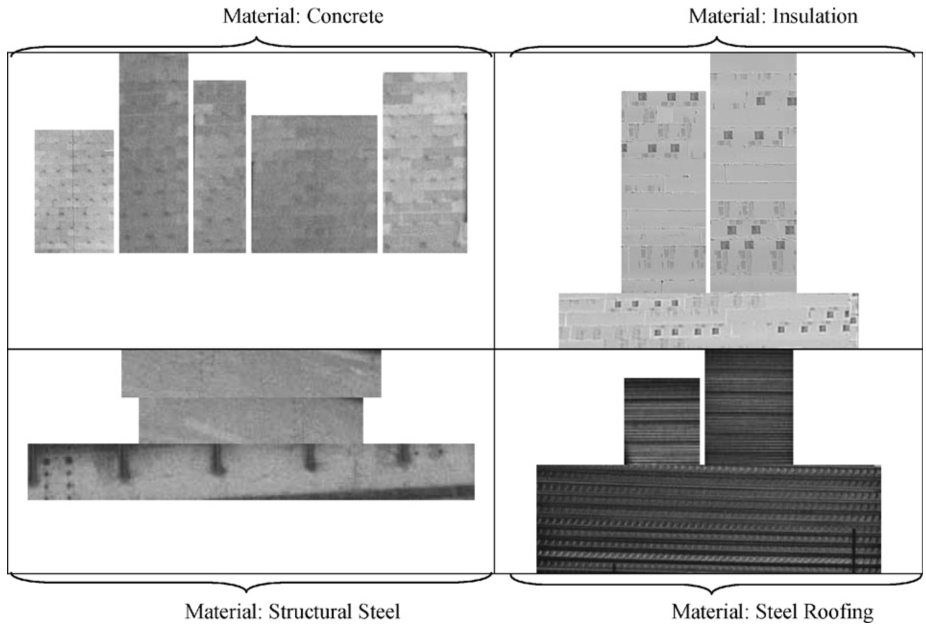

Fig. 13. Knowledge database example.

Table 1
Example signature values of concrete wall sample

Concrete_wall0.tif	Mean	Variance
Intensity	0.664516	0.00219
Normalized red	0.332387	5.21E-05
Normalized green	0.334969	1.31E-05
Normalized blue	0.322643	4.95E-05
Spot Texture 1	0.828271	0.010283
Spot Texture 2	0.185624	0.009647
Line Texture 1	0.510351	0.00502
Line Texture 2	0.496911	0.003676
Line Texture 3	0.828271	0.010283
Line Texture 4	0.185624	0.009647

with the existing samples reveals redundant sampling (over-sampling) and rejects the redundant information. New samples of an existing material that cannot be retrieved from the existing samples are kept as additional information. This way, the user doesn't have to manually decide the appropriateness or redundancy of new samples.

The variety of materials that can be detected with this method depends on the variety of samples that exist in the sample database. Subcontractors for example, would most likely need a smaller variety of material samples than general contractors and facility owners. In other words, there is no perfect set of samples that can cover all the needs of each project participant, which justifies the flexibility of this method in formulating the knowledge database. Materials with varied appearances (e.g., painted differently or wet vs. cured concrete) can also be considered by adding more samples in the database to account for those differences.

In order to identify the materials within the given image, each cluster signature is then compared to those in the knowledge database, and whenever a match is found (as defined below) the cluster inherits the material name of the sample it was matched to. This process is fast since the knowledge database signatures are pre-computed and is crucial in determining the precision and recall of the overall method, since accurate material recognition is the key element. For this reason, the comparison is performed in two stages by using:

(1) Thresholds: It is necessary to define thresholds for each signature element in order to constraint the space inside which different possible matches can occur and ensure the independency of each comparison. This is due to the fact that a distance function alone will not discard cases of low relevance and will instead simply rank them low. By utilizing a set of thresholds to define the limits for each feature, for a match to occur, most unrelated clusters are immediately disqualified and are not considered in the steps that follow.
(2) Distance: It is important to match image regions based on feature distance especially when materials with "similar" visual features are present (e.g., concrete blocks and grouting). In such cases, thresholds will discard most other possibilities, but a proper distance function will assist in selecting the best possible match among those who have qualified. The distance function itself was the Euclidean distance of the normalized difference of each signature value.

The resulting indexing scheme is static for each detected similarity. Images are indexed based on each material that exists in the knowledge base. The indexing results are stored in "image objects" that comprise of the original image along with its indexing results for each material sample. When new material samples are added, the user can direct the method to search for that material in all existing images and store the additional information (of existence/non-existence) in each image object.

An example is presented here in order to better familiarize the reader with this method. A site engineer notices a crack on a concrete beam and wishes to find when this crack first appeared. Since the date information is what the user is looking for, the last image (according to timestamp) that shows this beam with no crack is needed. The problem here is to find all the images that depict this beam from the thousands of images in the database, so that their ranking according to time and subsequent visual review can reveal the time when the crack first appeared. The engineer seeks for images with concrete using this method. The missing information is the exact date that the crack appeared. Since all images in the database have been pre-classified according to the materials contained in them, the method initially loads the database of images and searches for concrete in their signatures. Following that, it discards images that don't have any concrete and ranks the remaining images according to the amount of concrete detected inside. The ranked images are then presented to the user. The engineer then browses through those images to find the oldest picture that shows the desired beam with the undesired crack.

3.2. Results and validation

To validate the proposed methodology several tests were conducted on an image collection of 1025 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 size is in the range of [2; 377]. The groups were defined (by an independent researcher to avoid bias) based on the Masterformat [24] standards as follows: (a) material site construction: wood (225 images), gravel (13 images), and earth (377 images), (b) concrete: cast-in-place concrete (279 images), scaffolds (105 images), forms (199 images), and rebar (267 images), (c) masonry: concrete blocks (17 images), (d) metal: steel (156 images), metal deck (27 images), metal framing (27 images), and metal studs (2 images), (e) thermal and water insulation: water (10 images), (f) finishes: drywall (7 images), paint (196 images), ceramic (40 images), (g) mechanical: pipes (64 images), exhaust (6 images), (h) electrical: conduits (11 images), smooth conduits (6 images).

From these groups, the following three were chosen for further testing for statistical significance purposes: earth, concrete, and paint. Following that, seven 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 cropped into regions and the clustering results are shown in Table 2.

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 (3000+) 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. Finally it is worthwhile noting that several clusters in most images contained more than 100,000 pixels and increased the accuracy and speed of the implemented solution significantly.

Table 2
Clustering statistics

	Average	Variance	Std. Dev.	Min	Max	Max–min
Avg. pixels	3479	4.6e + 6	2144	623	17,623	17,000
Pixels found	77%	1.6%	13%	18%	96%	78%

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 seven queries, once for each material group. The metrics used to assess the performance of this test are the standard precision and recall metrics; precision is the ratio of relevant images retrieved over all retrieved images, and recall is the ratio of relevant images retrieved over all relevant images. The results are depicted in standards precision-recall graphs, where the desired curve should resemble a one sided Laplacian curve (high precision at low recall; lower precision at higher recall).

The precision values were then adjusted to reflect the differences in generality (images in material group/all images) among the materials by multiplying each result with the ratio of generality/average generality: adj. precision $(g) = \text{Precision} * g/g_{\text{avg.}}$ This way, materials that appear most frequently (high generality e.g., earth, since most construction site images contain earth) do not artificially boost the overall results and materials that are rare (low generality) do not artificially lower them. It is important to note that the tolerances for each feature comparison were set based on the average tolerance of each feature for each material so as to increase the quality of the results.

Overall, similarly to the majority of retrieval mechanisms, the performance of this method lies in the successful ranking of relevant items retrieved. In this case, successful ranking is achieved when the precision at low recall (e.g., the top 20 relevant images) is higher when compared to precision at higher recall and preferably reaching 100%. The results (samples in Figs. 14–16) indicate that the preci-

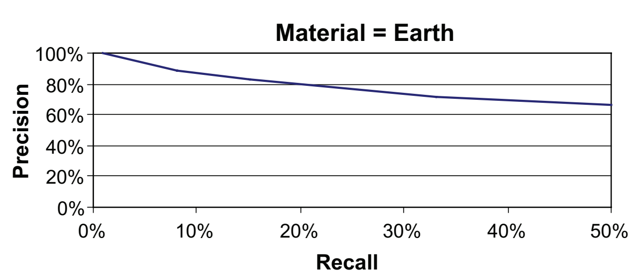

Fig. 14. Precision recall graph for earth.

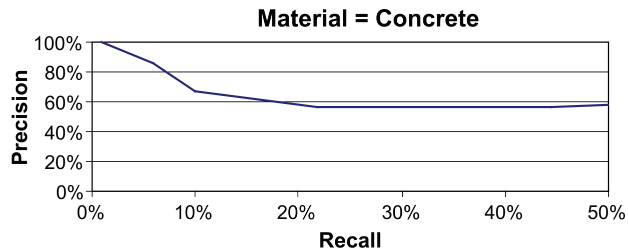

Fig. 15. Precision recall graph for concrete.

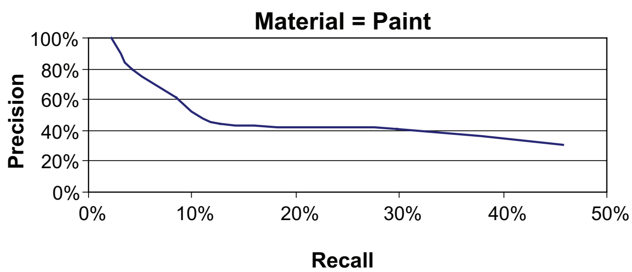

Fig. 16. Precision recall graph for paint.

sion at varying recall percentages is high, and at low recall it can reach 100%. In these figures it is also visible that feature-rich materials, such as earth, that take advantage of the entire feature spectrum (color, texture, etc.) are better recognized than materials with fewer distinctive characteristics such as paint (where only the color is distinctive).

4. Conclusions

The large amount of pictures collected daily in construction sites and the time needed to manually classify it motivated the researchers to investigate methods for automatically exploring the content of pictorial data in search for patterns that are interesting from a construction manager's perspective. The main objective was to develop novel criteria for the classification and retrieval of construction site images. Specifically, the goal of the classification step was devise methods for the meaningful and automated indexing of pictorial data, while the goal for the retrieval step was to find images based on criteria that the user is familiar with. Both of these goals were accomplished with the development of the material cluster recognition approach.

This paper presented a novel material based image retrieval method that is based on the recognition of material clusters in each image. Under this method, the pixels of each image are grouped into meaningful clusters and are subsequently matched with a variety of pre-classified material samples. This way, the existence (or not) of construction materials in each image is detected and later used for image retrieval purposes. This method allows engineers to search for construction images based on their content in a meaningful manner.

The evaluation of the results has shown that the Materials Identification method for the retrieval of construction site images can successfully extract material information from each image and allow for the comparison of material signatures instead of image signatures. It retains and enhances the advantage of user-friendliness of the Generic CBIR approach by using concepts that construction engineers are more familiar with 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 existing image retrieval methods as explained in Section 3 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] I. Abdel-Qader, O. Abudayyeh, M.E. Kelly, Analysis of edge-detection techniques for crack identification in bridges, J. Comp. Civ. Eng. 17 (4) (2003) 255–263.
[2] D. Koller, K. Daniilidis, H.H. Nagel, Model-based object tracking in monocular image sequences of road traffic scenes, Int. J. Comp. Vision 10 (3) (1993) 257–281.
[3] A.H. Aydilek, S.H. Oguz, T.B. Edil, Digital image analysis to determine pore opening size distribution of nonwoven geotextiles, J. Comput. Civ. Eng. ASCE 16 (4) (2002) 280–290.
[4] E.J. Garboczi, Three-dimensional mathematical analysis of particle shape using X-ray tomography and spherical harmonics: Application to aggregates used in concrete, Cem. Concr. Res. 32 (10) (2002) 1621–1638.
[5] D.F. Lin, H.Y. Wang, H.L. Luo, Assessment of fire-damaged mortar using digital image process, J. Mat. Civ. Eng. 16 (4) (2004) 383–386.
[6] A.H. Aydilek, M. Guler, Use of image analysis in determination of strain distribution in geosynthetic tensile testing, J. Comput. Civ. Eng. ASCE 18 (1) (2004) 65–74.
[7] A.M. Hartman, M.D. Gilchrist, Evaluating four-point bend fatigue of asphalt mix using image analysis, J. Mat. Civ. Eng. 16 (1) (2004) 60–68.
[8] M. Muste, Z. Xiong, J. Schone, Z. Li, Validation and extension of image velocimetry capabilities for flow diagnostics in hydraulic modeling, J. Hydraulic Eng. 130 (3) (2004) 175–185.
[9] P.H. Chen, Y.C. Chang, L.M. Chang, P.C. Doerschuk, Application of multi-resolution pattern classification to steel bridge coating assessment, J. Comput. Civ. Eng. 16 (4) (2002) 244–251.
[10] M.Q. Feng, C. Liu, M. Shinozuka, Electromagnetic image reconstruction for damage detection, ASCE J. Eng. Mech. 126 (7) (2000) 725–729.
[11] A. Bovik, Handbook of Image and Video Processing, first ed., Academic Press, 2000, ISBN 0-12-119790-5.
[12] N. Weyrich, G.T. Warhola, Wavelet shrinkage and generalized validation for image denoising, IEEE Trans. Image Process. 7 (1998) 82–90.
[13] D. Forsyth, J. Ponce, Computer Vision – A Modern Approach, first ed., Prentice Hall, 2002, ISBN 0130851981.
[14] A.C. Bovik, T.S. Huang, D.C. Munson Jr., The effect of median filtering on edge estimation and detection, IEEE Trans. Pattern Anal. Mach. Intell. PAMI 9 (1987) 181–194.
[15] D. Marr, E. Hildreth, Theory of edge detection, Proc. Roy. Soc. Lond. B 270 (1980) 187–217.
[16] J. Canny, A computational approach to edge detection, IEEE Trans. Pattern Anal. Mach. Intell. PAMI 8 (1986) 679–698.

[17] F.A. Velasco, J.L. Marroquin, Growing snakes: active contours for complex topologies. Elsevier Science, Pattern Recognition 36 (2003) pp. 475–482.
[18] I. Brilakis, L. Soibelman, 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 (2005).
[19] O.Y. Abudayyeh, Audio/Visual information in construction project control, J. Adv. Eng. Softw. 28 (2) (1997).
[20] B. Kosovac, T. Froese, D. Vanier, Integrating heterogeneous data representations in model-based AEC/FM systems. In: G. Gudnason (Ed.), Proceedings of Construction Information Technology 2000 – CIB-W78, IABSE, EG-SEA-AI International Conference on Information Technology, Reykjavik, Iceland, 28–30 June, 2000, Icelandic Building Research Institute, vol. 2, 2000, pp. 556–567.
[21] J. Malik, P. Perona, Preattentive texture discrimination with early visual mechanisms, J. Opt. Soc. Amer. 3 (10) (1990) 1673–1683.
[22] S.G. Krantz, Handbook of Complex Variables, Birkhäuser, Boston, MA, 1999, p. 16.
[23] S.W. Smith, The Scientist and Engineer's Guide to Digital Signal Processing, California Technical Publishing, 1997, ISBN 0-9660176-3-3.
[24] MasterFormat^{TM} 2004 Edition, Construction Specifications Institute (CSI) and Construction Specifications Canada (CSC) ISBN 0-9762399-0-6, Paperback.